Energetics and dynamics of lipase A activation by the steric chaperone Lif in *Pseudomonas aeruginosa*

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Neha Verma

aus

Indien

Düsseldorf, August 2019

Aus dem Institut für Medizinische Chemie und Pharmazie

der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Holger Gohlke

Korreferent: Prof. Dr. Karl-Erich Jaeger

Tag der mündlichen Prüfung: 26.08.2019

Eidesstattliche Erklärung

lch versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Diese Dissertation wurde in der vorgelegten oder einer ähnlichen Form noch bei keiner anderen Institution eingereicht und es wurden bisher keine erfolglosen Promotionsversuche von mir unternommen.

Düsseldorf, im August 2019

In loving memory of my grandfather

“Veer tum bade chalo, Dheer tum bade chalo”

TABLE OF CONTENTS

LIST OF PUBLICATIONS

This thesis is based on the following publications:

I. Jakub Kubiak[§], **Neha Verma[§]**, Peter Dollinger, Holger Gohlke, Claus A.M. Seidel, Filip Kovacic, Karl-Erich Jaeger

Functional dynamics of a structurally minimalistic chaperone

In preparation 2019.

II. Aldino Viegas[§], Peter Dollinger[§], **Neha Verma[§]**, Jakub Kubiak, Thibault Viennet, Claus A.M. Seidel, Holger Gohlke, Manuel Etzkorn, Filip Kovacic, and Karl-Erich Jaeger

Structural and dynamic insights revealing how lipase binding domain MD1 of Pseudomonas aeruginosa foldase affects lipase activation

Submitted 2019.

III. **Neha Verma**, Peter Dollinger, Filip Kovacic, Karl-Erich Jaeger, and Holger Gohlke

The membrane-integrated steric chaperone Lif facilitates active site opening of Pseudomonas aeruginosa lipase A

Submitted 2019.

[§] These authors contributed equally to this work.

ABBREVIATIONS

αLP	Alpha-lytic protease
AHC	Average alpha-helical content
ATP	Adenosine triphosphate
AV	Accessible volume
***Bg*LipA**	*Burkholderia glumae* lipase A
***Bg*LipA$_c$**	Closed state of lipase A from *B. glumae*
***Bg*Lif**	Lipase-specific foldase from *B. glumae*
***Bg*Lif:LipA**	Lipase:foldase complex from *B. glumae*
CD	Circular dichroism
CNA	Constraint network analysis
COM	Center of mass
***D*$_{COM}$**	Distance between centers of mass
EHD	Extended helical domain
FF	Force field
FRET	Förster resonance energy transfer
ΔG	Change in Gibbs free energy
GSP	General secretion pathway
HB	Hydrogen bond
HSP	Heat shock protein
IM	Inner membrane
K$_D$	Dissociation constant
MD	Molecular dynamics
MD1	Mini-domain 1
MD2	Mini-domain 2

MEM	Maximum entropy method
NMR	Nuclear magnetic resonance
OCP	Octyl-phosphinic acid 1,2-bis-octylcarbamoyloxy-ethyl ester
OM	Outer membrane
***Pa*LipA**	*Pseudomonas aeruginosa* lipase A
***Pa*LipA$_c$**	Closed state of lipase A from *P. aeruginosa*
***Pa*LipA$_o$**	Open state of lipase A from *P. aeruginosa*
***Pa*Lif**	Lipase-specific foldase from *P. aeruginosa*
***Pa*Lif:LipA**	Lipase:foldase complex from *P. aeruginosa*
PDB	Protein data bank
PMF	Potential of mean force
$p(R_{DA})$	Probability of mean inter-dye distance distributions
R_{DA}	Inter-dye distance
RMSD	Root mean square deviation
SASA	Solvent accessible surface area
SB	Salt bridge
SEM	Standard error of mean
T2SS	Type-2 secretion system
TMD	Transmembrane domain
US	Umbrella sampling
VD	Variable domain
WHAM	Weighted histogram analysis method

ZUSAMMENFASSUNG

Bakterielle Lipasen stellen eine wichtige Gruppe von Enzymen im biotechnologischen und biomedizinischen Bereich dar. Aufgrund ihrer Fähigkeit Glycerinester zu hydrolysieren, sind sie in der Lebensmittel-, Waschmittel- und Pharmaindustrie sehr gefragt. In den letzten Jahrzehnten ist sowohl die biotechnologische Optimierung dieser Lipasen, als auch ihre biotechnologische Produktion auf die Bedürfnisse der Industrie angepasst worden und somit auch zu einem wichtigen Forschungsthema in der Wissenschaft geworden. Der erste Schritt der biotechnologischen Produktion der Lipasen ist die Überexpression, die sich allerdings als sehr schwierig herausstellt hat. So benötigen Lipasen aus *Pseudomonas* und *Burkholderia* die Unterstützung der Lipase-spezifischen Foldase Lif, damit sie ihre katalytisch aktive Konformation erreichen. Nur Lipasen in ihrer aktiven Konformation können industriell genutzt werden, sodass ein starker Bedarf besteht den Faltungsmechanismus dieser Lipasen auf molekularer Ebene zu untersuchen und zu verstehen.

Lipasen werden nach ihren Aminosäuresequenzen und biologischen Eigenschaften in acht Familien eingeteilt. Familie I umfasst die Vertreter der “echten” Lipasen, zu denen auch die Lipase A aus *Pseudomonas aeruginosa* (*Pa*LipA) zählt. Abgesehen von ihrer kommerziellen Relevanz ist *Pa*LipA ein bekannter Virulenzfaktor mit Relevanz in verschiedenen Krankheiten, wie Mukoviszidose. Ziel dieser Dissertation war es, den Aktivierungsmechanismus von *Pa*LipA durch die Lipase-spezifische Foldase *Pa*Lif mittels computerbasierten Methoden, zusammen mit einer strukturellen Charakterisierung durch Fluoreszenzmessungen und *in vitro* Experimenten, zu untersuchen.

Zunächst wurde eine computerbasierte Untersuchung der Foldase *Pa*Lif in Abwesenheit der entsprechenden Lipase durchgeführt. Dies ermöglichte die Charakterisierung der Funktionsdynamik der einzelnen Domänen von *Pa*Lif auf atomarer Ebene. Zweitens wurden NMR-Strukturen für die stabile Domäne MD1 von *Pa*Lif gelöst und die strukturellen

Determinanten der MD1 Domäne bei der Aktivierung von *Pa*LipA untersucht. Basierend auf diesen Ergebnissen wurde ein potentieller Aktivierungsmechanismus abgeleitet, in dem auch weitreichende Wechselwirkungen, die sich von *Pa*Lif bis zum aktiven Zentrum von *Pa*LipA erstrecken, eine Rolle spielen. Drittens wurde die Öffnungsenergie des aktiven Zentrums von *Pa*LipA in An- und Abwesenheit von *Pa*Lif durch Molekulardynamiksimulationen quantifiziert. Zudem analysierte ich die wichtigsten Wechselwirkungen zwischen *Pa*Lif und *Pa*LipA hinsichtlich der mechanischen Stabilität. Die vorliegende Arbeit liefert somit molekulare Einblicke in die Aktivierungsmechanismen von *Pa*LipA durch *Pa*Lif, was nützlich sein kann, um die biotechnologische Produktion von *Pa*LipA im großen Maßstab zu optimieren.

ABSTRACT

Bacterial lipases are an essential group of enzymes in the biotechnological and biomedical field. Due to their ability to hydrolyze esters of glycerol, they are highly demanded in food, dairy, detergent, and pharmaceutical industries. Over the last decades, optimization of these enzymes and their production to meet industrial needs has become a relevant research topic for the scientific community. For the production of these enzymes, the first step is overexpression, which turns out to be a difficult step. Lipases from *Pseudomonas* and *Burkholderia* need the assistance of the lipase-specific foldase Lif to achieve their active conformation before they are secreted and can be harvested. Therefore, there is a strong need to comprehend the folding mechanism of these lipases at the molecular level.

Lipases are classified according to their amino acid sequences and biological properties into eight families. Family I represents “true” lipases. The lipase A from *P. aeruginosa* (*Pa*LipA) is a prominent member of this family. Beyond its commercial relevance, it is a known virulence factor in many diseases such as cystic fibrosis.

The goal of this thesis was to investigate the activation mechanism of *Pa*LipA assisted by lipase-specific foldase *Pa*Lif by means of computational methods together with structural characterization by fluorescence measurements and *in vitro* experiments.

First, a computational investigation of *Pa*Lif in the absence of its cognate lipase was performed. This allowed the characterization of functional dynamics of the individual domains of *Pa*Lif at an atomic level. Second, solution nuclear magnetic resonance structures have been solved for the structurally stable domain MD1 of *Pa*Lif. Based on this data, the structural determinants of this domain in the activation of *Pa*LipA were investigated. These findings provide a putative mechanism of long-range interactions spanning from *Pa*Lif to the active site of *Pa*LipA. Third, the energetics of active site opening in *Pa*LipA with and without its foldase is quantified by

unbiased and biased molecular dynamics simulations. Finally, I analyzed the key interactions between *Pa*Lif and *Pa*LipA with respect to mechanical stability, which provides the basis for a possible mechanism of *Pa*LipA activation via *Pa*Lif. The work done in this thesis provides a detailed understanding of the mechanistic insights of the activation process of *Pa*LipA, which can be useful to optimize the production of this enzyme on a large scale.

1. INTRODUCTION

Over a century ago, C. Eijkman reported for the first time that several bacteria produce and secrete lipases [1]. Lipases have an unusual catalytic activity to hydrolyze and synthesize long-chain acylglycerols [2] with high chemo, regio- and stereoselectivity for their substrates [1]. Hence, they constitute the most important and widely used class of biocatalyst for many industrial, pharmaceutical, and biotechnological applications [1] (**section 2.3.1**). The major limitation of the use of lipases at the industrial scale is their high production cost. Therefore, many studies have been initiated in order to optimize their production and to develop these enzymes into the perfect tool for biotechnology The production of these enzymes is a challenging task, as it is a multi-step process in which approximately 30 other intracellular proteins are involved to ensure correct folding and activation of the respective lipase (**section 2.4.1**) [1]. In addition, lipases require the complex type-2 secretion system (T2SS) (**section 2.2.2**) in order to be secreted to extracellular medium [3]. Until now, the detailed molecular mechanism of lipase folding and secretion is still unknown, which became the motivation for this thesis.

Bacterial lipases are classified into eight different families according to their fundamental biological properties and amino acid sequences [4]. Lipases from *Pseudomonas* and *Burkholderia* belong to the subfamily I.1 and I.2 of the family I, respectively. Lipase A produced by the gram-negative bacterium *Pseudomonas aeruginosa* (*Pa*LipA) came into the spotlight over a past decade due to its remarkable catalytic properties and its role as a virulence factor in many diseases (**section 2.1 and 2.3**) [5, 6]. It is one of the most prominent and well-studied proteins of the “true” lipase family [7] with a solved structure and also the main subject of this thesis. *Pa*LipA is synthesized in the cytoplasm and translocated to the periplasm via Sec translocase (**section 2.2**) [8], where it folds spontaneously to an inactive, near-native (intermediate) state. This near-native state of *Pa*LipA requires the assistance of the specific

foldase *Pa*Lif for its conversion into an enzymatically active conformation, which is the critical step of *Pa*LipA biogenesis (Figure 1) [9]. The requirement of such a foldase indicates a high energy folding barrier between the near-native and native state of *Pa*LipA, and *Pa*Lif helps *Pa*LipA to cross this barrier (**section 2.4.2)**. Once activated, *Pa*LipA traverses the bacterial outer membrane by using the T2SS-involving Xcp machinery (**section 2.2**) [10].

Until now, the crystal structure of many homologous lipases have been solved, and they all exhibit a unique core structure known as the α/β-hydrolase fold (**section 2.4**) [11-13]. All enzymes from this family depend on the presence of three amino acids to form a catalytic triad, which plays key roles in the catalytic reaction. Usually, a nucleophilic serine, a histidine, and an aspartate or a glutamic acid form the catalytic triad at the binding site of lipases [14]. Lipases also contain an oxyanion hole and a lid domain [15]. Remarkably, helix α5 forms the lid, which shields the catalytic triad residues from the solvent, and thus prevents the substrate from entering the binding site [16]. Several lipases are crystallized in the closed inactive conformation where the lid domain occludes the binding site (**section 2.4**). The crystal structure of *Pa*LipA (PDB code: 1EX9) has been crystallized in an open conformation in the presence of inhibitor, whereby helix α5 is displaced from the active site [17], and an additional partial β-sheet structure has formed in the region of residues 17-30, which forms the cleft of the binding site (**section 2.4**). Despite the presence of the lid, lipases from *P. aeruginosa* and *B. glumae* are exceptions as they do not show interfacial activation [18]. The crystal structure of lipase A from *B. glumae* (*Bg*LipA) in complex with its specific foldase (*Bg*Lif:LipA) (PDB code: 2ES4) [19] was found to be virtually identical to that of free inactive closed lipase A (BgLipA$_c$) (PDB code: 1QGE) [15]. In the complex, helix α5 is in the closed conformation while residues 17-30 formed a partial β-sheet structure, which is a characteristic feature of open lipases (**section 2.4**). Furthermore, the addition of substrate to the crystals results in lipase activity without affecting the crystal integrity [19], which indicates that *Bg*LipA in the complex is like a “loaded spring”

ready to move to the open state. It strongly suggests that the *Bg*LipA is in an intermediate state in the complex on the way to an open state. Here, I applied the aforementioned computational methods (**section 2.6**) in combination with experimental data to investigate the above stated hypothesis and the activation process of *Pa*LipA facilitated by *Pa*Lif.

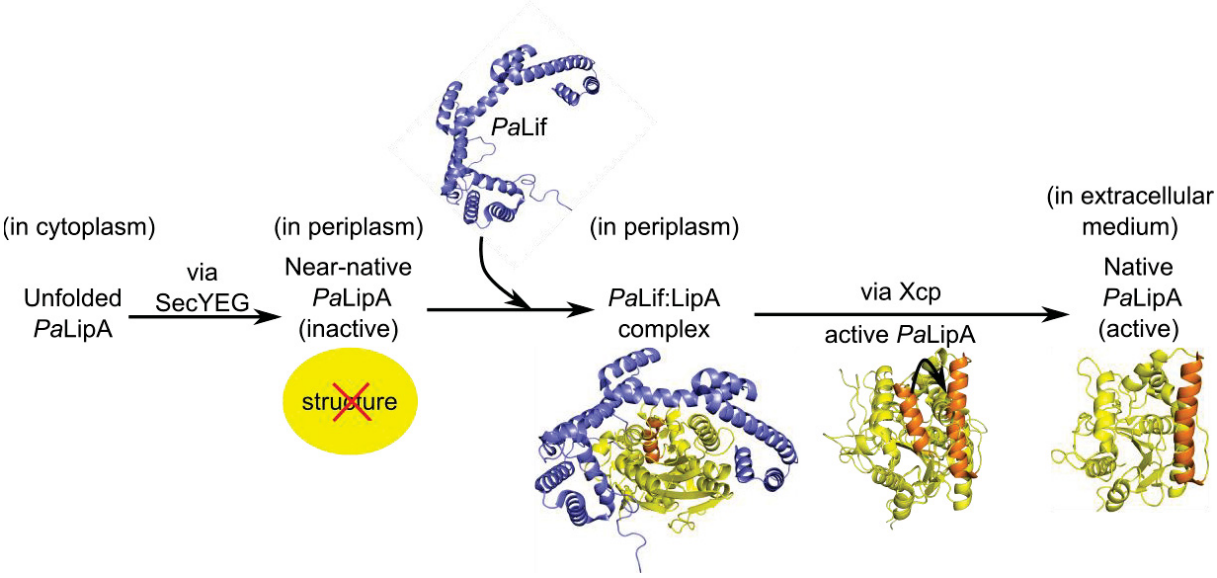


Figure 1. Proposed hypothetical folding model for *Pa*LipA. Unfolded *Pa*LipA is translocated from the cytoplasm to the periplasm via SecYEG transporter. It was observed that in the absence of *Pa*Lif, *Pa*LipA folds to a near-native state, inactive state. Upon addition of *Pa*Lif, this near-native, intermediate state of *Pa*LipA forms a stable complex with the foldase (*Pa*Lif:LipA). In the complex, *Pa*LipA undergoes conformational changes to fold into the active native state and exports out to the extracellular medium by the Xcp machinery of T2SS. This hypothetical model is based on *in vitro* experiments and available crystal structures. However, a detailed molecular mechanism of *Pa*LipA activation via *Pa*Lif remains elusive. Figure adapted from ref. [20].

Furthermore, in order to understand the *Pa*LipA activation mediated by *Pa*Lif in atomic detail, it was essential to study the structural and functional dynamics of *Pa*Lif first. So far, no X-ray structure is available for free *Pa*Lif, and very little is known about its mechanism of action in general. *Pa*Lif belongs to a small class of steric chaperones. In the crystal structure of *Bg*Lif:LipA complex, *Bg*Lif was found to form a unique "head-phone" like fold, with two mini-domains (MD1 and MD2) connected by an extended helical domain (EHD).

Here, to understand the conformational dynamics of *Pa*Lif in the free state, an integrative approach combining molecular dynamics (MD) simulations and Förster Resonance Energy Transfer (FRET) experiments (**section 2.6**) was applied. This resulted in an ensemble of *Pa*Lif

structures ranging from compact to extended conformations, which increases the resolution of FRET data beyond shape information to atomistically detailed structures of minimalistic structured *Pa*Lif [21]. Based on the results obtained from the above study, solution structures of MD1 of *Pa*Lif were also solved and investigated for its role in *Pa*LipA activation.

Overall, my work sheds light on the underlying mechanism of *Pa*LipA folding from an inactive to an active state, catalyzed by the steric chaperone *Pa*Lif. Furthermore, it provided insights into the functional dynamics of structurally minimalistic *Pa*Lif. This work will enable us to understand the mechanism of action of steric chaperones in more detail and provide relevant insights to unravel further mechanistic understandings of this interesting system. This can be of particular interest for biotech companies for a cost-effective overproduction of lipase.

2. BACKGROUND

2.1 The microorganism *Pseudomonas aeruginosa*

In the 19th century, Migula identified the *Pseudomonas* species for the first time as anaerobic, gram-negative bacilli measuring 0.5 to 0.8 μm (diameter) by 1.5 to 3.0 μm (length) [6] (Figure 2). The *Pseudomonas* genus pool contains more than 140 species, of which more than 25 species are associated with humans [22]. Approximately 80% of Pseudomonas infections are caused by *P. aeruginosa* and *P. maltophilia* [22]. The species name aeruginosa originated from the Latin word "Verdigris", which means copper rust, referring to its distinctive blue-green color on solid media. This blue-green color is due to the water-soluble pigments pyocyanin, and pyoverdine produced by *P. aeruginosa* [23].

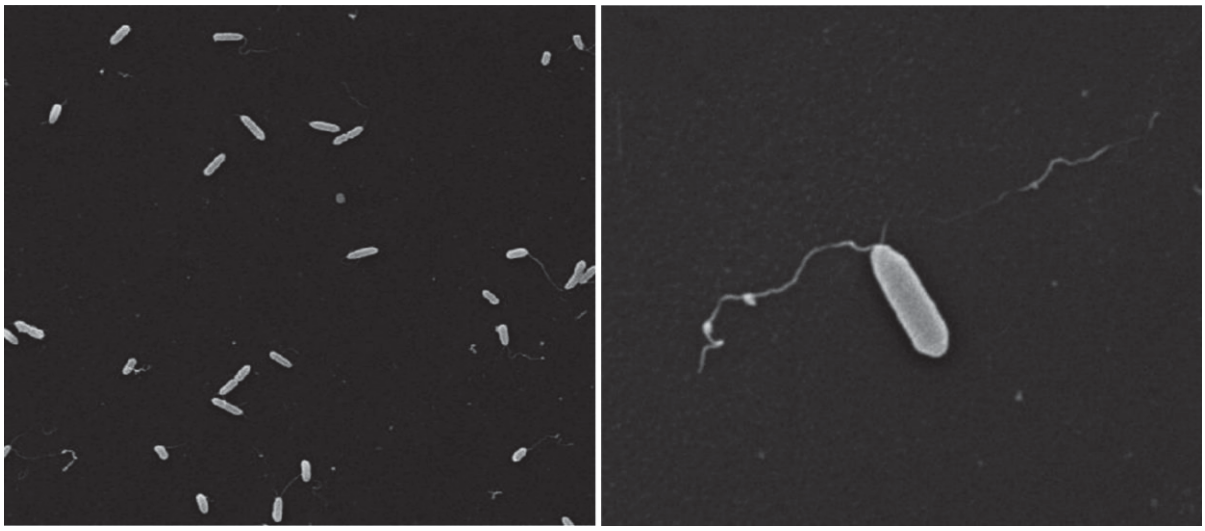


Figure 2. Micrographic image of gram-negative bacterium *P. aeruginosa*. Scanning electron micrographic image of *P. aeruginosa* bacilli adherent to the collagen surface (left), a high power scanning micrographic image of same bacilli with polar pili along the bacterial long axis (right). Images are adapted from [27].

The genome of *P. aeruginosa* is vast, complex, and consists of a single circular chromosome. The complete sequence of the first *P. aeruginosa* genome was reported in 2000 for strain PAO1 [24]. It has a relatively large genome of the size of 6.3 million base pairs with high G + C content (65-67%) and contains 5,570 predicted open reading frames [25]. This genome encodes for a wide variety of enzymes, which are involved in various metabolic pathways and transport

systems [26]. Specifically, the number of genes predicted to encode for outer membrane proteins is disproportionally large (~ 150) as compared to inner membrane proteins. Also, approx. 8.4% of the genome encodes regulatory genes, which allows the bacterium to adjust to ubiquitous growth environment [24].

P. aeruginosa produces many virulence factors which not only affect plants and animals but also cause severe infections in humans [28]. *P. aeruginosa* may grow for decades in the lungs of an individual with cystic fibrosis. Cystic fibrosis transmembrane conductance regulator inhibitory factor, produced by *P. aeruginosa,* indirectly stimulates neutrophils and promotes the sustained infection in patients [29]. The major virulence factors of *P. aeruginosa* are the type III secretion effector proteins exotoxin S, T, U, and Y. ExoS and T both are bifunctional toxins [30]. They act by disrupting the host cell signaling pathway required to stimulate the phagocytic NADPH oxidase [29]. There are many more well-characterized virulence factors produced by *P. aeruginosa*; comprehensive information can be found in refs. [30-34].

P. aeruginosa is regarded as a promising microorganism in various biotechnological applications, industries, and commercial sectors. It is an excellent source of enzymes acting as catalysts in specific biochemical reactions [35]. High specificity and efficiency make them a valuable catalyst in many biotechnological industries, including pharmaceutical, food, cosmetics, detergents, paper, and pulp industries [35]. Furthermore, *P. aeruginosa* produces less toxic and eco-friendly biosurfactants and polycyclic aromatic hydrocarbons (PAHs) [36]. Their emulsification activity against different hydrocarbons made them an essential agent in bioremediation [37].

In conclusion, *P. aeruginosa* is not only a source of virulence factors but also has wide applications in various industries, which made it one of the most studied gram-negative bacteria.

2.2 Protein export and secretion in gram-negative bacteria

2.2.1 Cell envelop of gram-negative bacteria

Gram-negative bacteria have evolved a sophisticated and complex cell envelop in order to survive in an unpredictable and hostile environment (Figure 3) [38, 39]. The cell envelop has different functions such as protecting the cell in a hostile environment, maintaining shape and stability and allowing communication with the environment to allow selective passageways of nutrients from the outside and waste products from the inside [39].

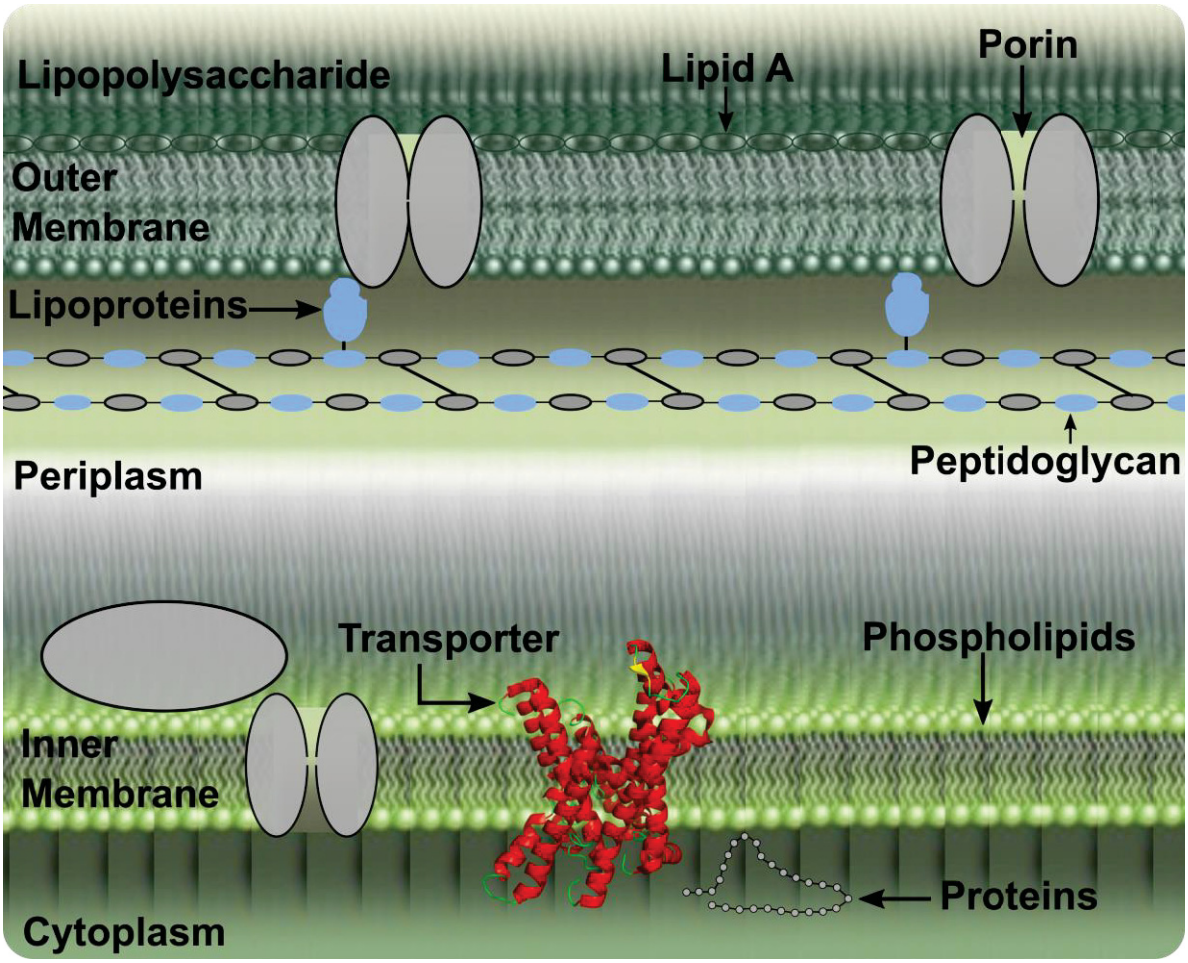


Figure 3. Schematic representation of the cell envelop of gram-negative bacteria. Figure modified from Douzi *et al*. (2004) [41].

Gram-negative bacteria, compared to gram-positive ones, have an additional outer membrane (OM) located above a thin layer of peptidoglycan, in the outer leaflet (extracellular side). It is an asymmetric lipid bilayer with selective permeability, made of lipopolysaccharides (LPS) and

phospholipids [38]. A large number of integral pore proteins and porins allow the free passage of small hydrophilic molecules and water across the membrane. Between the OM and inner membrane (IM), a narrow aqueous concentrated gel-like matrix called periplasm is found (Figure 3) [39].

IM mainly consists of phosphatidylethanolamine and phosphatidylglycerol and plays a major role in the synthesis of phospholipids, peptidoglycan, and lipopolysaccharides, via polysomes and membrane-bound proteins [40].

The periplasmic space is a multipurpose compartment which is more viscous than the cytoplasm and densely packed with proteins [39, 42]. It allows numerous mechanisms of protein oxidation, chemotaxis, chaperone-guided folding, regulation of cell division, efflux, and signaling [42, 43]. Despite all functions, the size of the periplasmic space has been highly debated in the past. Recently, Collins and co-workers found that the crystal structure of Wza - Wzc, the capsular transporter complex, is suspended into the periplasm from the inner leaflet of OM, which leads them to conclude that the periplasmic thickness should be of ~145 Å [44]. This value is also in agreement with the estimation of ~150 Å obtained from the structural model of the needle complex for the type III secretion pathway, resolved via cryoelectron microscopy [45, 46].

Usually, the periplasm contains autolysin, which along with penicillin-binding proteins, construct the peptidoglycan layer [47]. Peptidoglycan (murein sacculus) is made up of glycan strands composed of *N*-acetyl glucosamine and *N*-acetyl muramic acid, cross-linked by pentapeptides [48]. The OM is stitched to the peptidoglycan by a lipoprotein called murein lipoprotein. This lipoprotein has a lipid moiety attached to its amino terminus, which embeds it in the OM [49].

The inner leaflet of OM contains phospholipids, whereas the outer leaflet is composed of LPS molecules [50]. The OM proteins are mainly of two types: lipoproteins and β-barrel proteins

[39, 50]. The tight packing of acyl chains and LPS molecules as well as the porins activity make the OM an effective barrier for hydrophobic molecules [39].

2.2.2 Protein export and secretion

Gram-negative bacteria produce many proteins that are supposed to be exported to the periplasm or extracellular medium [51].

Protein export system

Approximately 20% of the proteins are synthesized outside the cytoplasm and need to be translocated across the cytoplasmic membrane [52]. The main pathway for the translocation of these proteins into the IM is the Sec pathway [53]. The Sec translocon is located in the cytoplasmic membrane and has three subunit cores, namely SecY, SecE, and SecG, together designated as SecYEG. In 2002, a three-dimensional reconstruction of the SecYEG was obtained by cryoelectron microscopy (PDB code: 3DIN) (Figure 4A) [54]. The Sec translocon translocates substrate across the membrane as well as also inserts them into the membrane laterally. All substrates of Sec translocase contain a hydrophobic N-terminal signal sequence, and the membrane proteins contain a membrane anchor signal for insertion [8].

A newly synthesized preprotein needs to be maintained in an unfolded state for export across the IM. SecB is a cytosolic molecular chaperone that binds to preproteins and maintains them in a translocation-competent state and prevents aggregation [55]. After forming the complex with the preprotein, SecB binds to SecA and transfers the preprotein to SecA for further translocation. SecA is an ATPase motor protein that provides the required energy for the translocation of the hydrophilic region of protein across the membrane (Figure 4A) [8]. Some preproteins are targeted to the translocon by signal recognition particle, and then they bind to SecA for translocation [8]. Unfolded preproteins are translocated by Sec translocase across the IM, whereas folded proteins are translocated by the twin-arginine translocase (Tat). Further details about the Tat pathway can be found in refs. [8, 56, 57].

Protein secretion system

In gram-negative bacteria, proteins can be secreted to the extracellular medium directly from the cytoplasm, which is a one-step mechanism [6]. In this, proteins can be secreted via a tunnel crossing the IM, periplasm, and OM. In other cases, proteins are first exported to the periplasm via SecYEG or TatABCE and then to the extracellular medium, also called two-step mechanism using transporter [51]. In order to cross two hydrophobic barriers, gram-negative bacteria have up to six distinct secretion systems. Of these, type 1, 3, 4, and 6 secretion systems follow a one-step secretion mechanism, which does not involve periplasmic intermediates, whereas type 2 and 5 use a two-step secretion system [51]. Comprehensive information about type 1, 3, 4, 5, and 6 can be found in refs. [51, 58-63].

Type 2 secretion system

T2SS is also known as the general secretion pathway (GSP) in gram-negative bacteria [51]. T2SS secretes a large number of enzymes, toxins, and virulence factors like exotoxin, cellulase, lipase, elastase, and alkaline phosphatase [64]. It is a two-step secretion pathway. Hence first the substrate is secreted to the periplasmic space from the cytoplasm via the Sec or Tat machinery, and in the second step, after the translocation in the periplasm, the secretion-competent protein state is secreted to extracellular space by the so-called GSP proteins or type 2 secretion machinery. The T2SS machinery is built up of 14 different GSP proteins, from which 12 GSP proteins are associated with the IM [65]. The 14 proteins form a sophisticated machinery with the ability to discriminate between the proteins that need to be transported across the OM and that reside in the periplasm [13].

In *P. aeruginosa*, the secretion machinery used in T2SS is called Xcp (Figure 4B), whereas another type 2 secretion system was also identified in *P. aeruginosa,* which is called Hxc-system [10, 66]. The Xcp machinery consists of 12 GSP proteins Xcp A and Xcp P to Z, which are organized into two operons [67]. The function of most Xcp proteins is still not entirely

known, but some components are well understood, like Xcp T-X, which act as pseudopilins [68]. These pseudopilus structures get inserted in the periplasm and act as a piston to push the substrate through the OM [69].

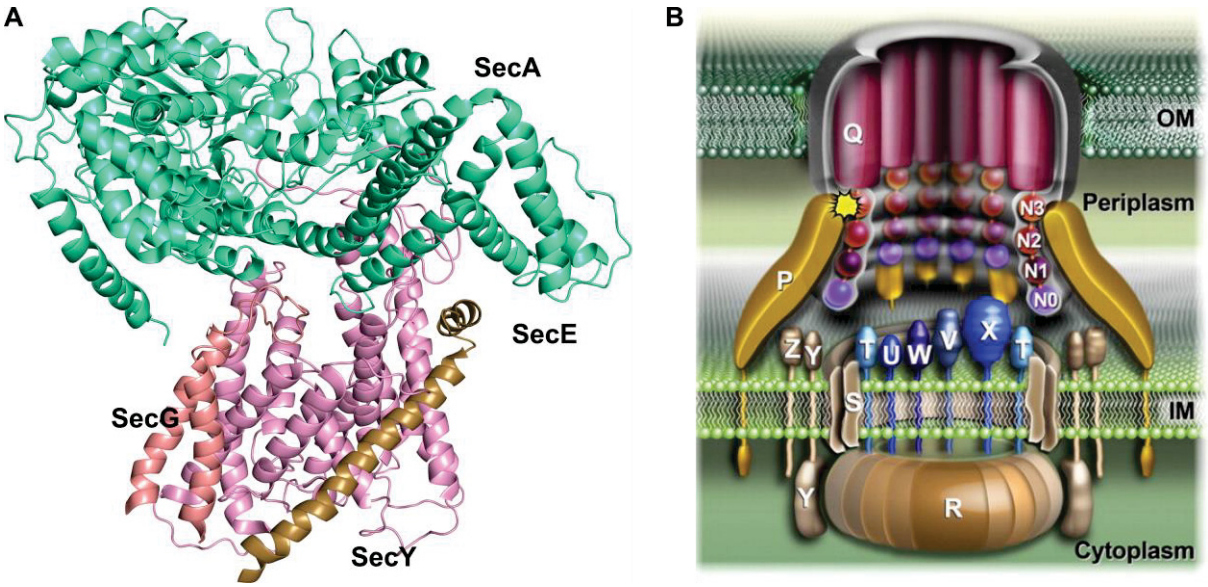


Figure 4. Import-export transporters for lipase in the cell of a gram-negative bacterium. (A) Crystal structure of SecYEG translocon in complex with SecA (PDB code: 3DIN) [70]. **(B)** Schematic model of *P. aeruginosa* Xcp transporter. Figure adapted from Douzi *et al.* (2011) [41].

Xcp R provides the required energy, Xcp P is involved in the secretion of exoproteins through the secretin channel [6, 71], and Xcp RSYZ with Xcp P form an IM subcomplex, which acts as a platform for the organization of the other parts of the secreton [67]. Xcp Q is the OM protein and contains a conserved C-terminal domain and four N-terminal domains, which interact with the periplasmic domain of Xcp P to form the secretin channel [72]. Last but not least, the peptidase Xcp A processes the signal peptide of the subunits of pseudopilin, and it also methylates the amino-terminal residues of mature proteins [73].

Interestingly, T2SS is capable of secreting many different and unrelated substrates. It was suggested that exoproteins interact with the periplasmic domain of Xcp Q and P as well as with the pseudopilus tip [41], which stimulates Xcp R to provide energy for the pseudopilus function to push the substrate through the secretin channel [74, 75]. The secretion motif is responsible for distinguishing proteins that translocate across the OM and that resides in the periplasm. In

general, the substrates tend to be folded, rich in β-strands, and have structural similarity with the secretin domain [75].

Despite experimental evidences and several crystal structures solved for exoproteins secreted by the Xcp machinery, the complete mechanism of T2SS is still unknown. Moreover, how the transporter recognizes the conformational signal to secret the substrate across the OM has not been resolved yet. As a GSP in gram-negative bacteria, it is capable of transporting a wide range of toxins and folded proteins, making it interesting for research and further studies [72].

2.3 Bacterial lipases

Lipases secreted by gram-negative bacteria have evolved to efficiently catalyze the hydrolysis of triacylglycerols to glycerol, diacylglycerols, monoacylglycerols, and free fatty acids (Figure 5) [76, 77]. In addition to hydrolysis, they can reverse this reaction, which is known as transesterification in non-aqueous media [78]. Since lipases remain enzymatically active in organic solvents and catalyze reactions with high regio- and enantioselectivity [79], these enzymes are very useful tools for organic chemists and valuable biocatalysts in many biotechnological applications [1].

Earlier, two criteria have been used to classify a lipolytic enzyme as a “true lipase” [4]. First, as the lipolytic reaction occurs at the lipid-water interface, the enzyme should be activated by the presence of a lipid-water interface, which is known as interfacial activation. Second, the protein should contain a “lid” domain, which covers the active site and moves away in order to open the active site on contact with the interface so the substrate can bind [4]. Later, however, lipases were defined more precisely as carboxylesterases as the above criteria were found to be unsuitable for classification [80]. Lipases do not require any cofactor for their action as they belong to the class of serine hydrolases [81]. Triacylglycerols, found in dietary fat and oils, are a significant source of metabolic energy compared to carbohydrates and proteins [82], therefore lipases are essential key players in catabolism throughout all living organisms.

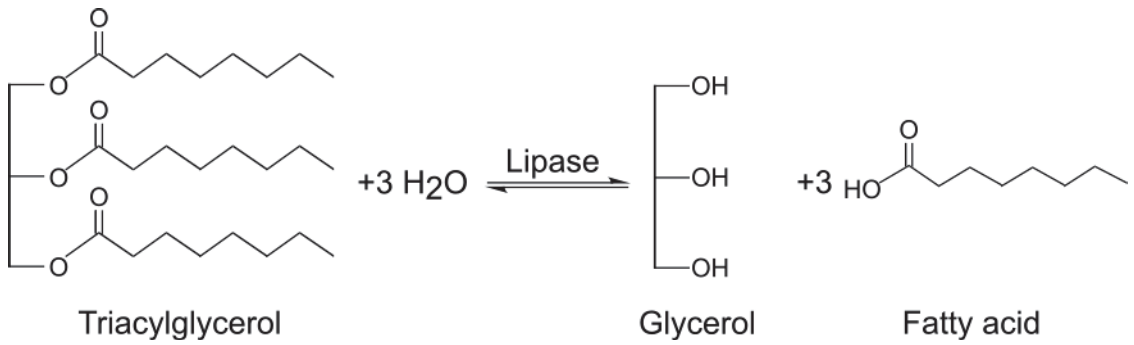


Figure 5. The hydrolysis reaction catalyzed by lipase enzyme. The hydrolysis reaction of triacylglycerol molecule to glycerol and free fatty acid, catalyzed by the lipase enzymes. This figure is taken from ref. [83].

Initially, bacterial lipases have been classified into eight families based on their amino acid sequence conservation and biological properties, and family I, which is further divided into six subfamilies, belongs to "true" lipases [4]. This broad classification was updated in 2002, and family I is divided into seven subfamilies [1]. The lipases secreted by *P. aeruginosa* belongs to subfamily I.1 with a molecular weight of approximately 30 kDa, and *B. glumae* belongs to subfamily I.2, which have a molecular weight of approximately 33 kDa [4]. Lipases from subfamily I.2 show about 60 % amino acid sequence identity with these of subfamily I.1 [4]. Additionally, they all contain a disulfide bridge, they need a specific foldase for correct folding, and they are secreted by T2SS (**section 2.2.2**) [84]. Bacterial lipases produced by *Pseudomonas* and *Burkholderia* species are the most studied lipases because of their ability to function in extreme conditions, their biotechnological potential and their relevance as crucial bacterial virulence factors (**section 2.1 and 2.3.1**) [85].

2.3.1 Commercial and biotechnological relevance of lipases

One of the major application of lipases in biotechnology is in polymer synthesis as they show high selectivity under mild reaction conditions. They catalyze the polymerization reactions for structurally complex monomers with multifunctional reactive groups [86]. Lipases with transesterification activity in organic solvents are considered the most critical biocatalysts in the production of biodiesel [87]. Lipases can convert vegetable oil to methyl- or other short

chain alcohol esters [1]. Lipases can catalyze enantioselective reactions for the synthesis of pharmaceuticals: *Pseudomonas* AK lipase was used to synthesize the chiral intermediate required for the synthesis of the antitumor agent epothilone A [88]. Lipases are extensively used in the resolution of racemic acids and alcohols [1]. They are also used for the production of herbicides, where only one of the enantiomers is active [1]. In cosmetics and flavors development, the most prominent example is the lipase-assisted synthesis of flavor and fragrance compounds such as (-)-menthol [89].

The first and vital use of lipases is in the detergent industry due to their unique property of catalysis at the water-lipid interface [1]. *B. glumae* lipase was patented for their detergent use by Unilever as it is very stable at different temperatures and shows catalysis under harsh washing conditions [90-92]. For the same properties, lipases are also used for the removal of fat-containing wastes in the paper industry, textile industry, and dye industry. The emulsification property of lipase brought them in the food industry, dairy industry, and baking industry [92]. They are used to produce low-calories fat, to emulsify oils, e.g., in the production of cocoa butter [93], for accelerating cheese ripening and flavor development. Lipases also substitute traditional emulsifiers used in baking, as the enzymes degrade polar wheat lipids to produce emulsifier *in situ* [94]. The high production costs are the most significant limitation of the industrial use of these enzymes nowadays.

2.4 Structure of *P. aeruginosa* lipase A

The first structure of bacterial lipase A was solved for *B. glumae* in 1993 by multiple isomorphous replacement techniques [95]. Later, *Chromobacterium viscosum* lipase A (*Cv*LipA$_c$) (PDB code: 1CVL) [15] and *B. glumae* lipase A (*Bg*LipA$_c$) (PDB code: 1QGE) [96] were crystallized in the closed conformation (Figure 6A). Structures of *Burkholderia cepacia* lipase A in complex with ligand R_c-(R_p,S_p)-1,2-dioctylcarbamoylglycero-3-*O*-*p*-nitrophenyl octylphosphonate (PDB code: 4LIP) [97] and without ligand (PDB code: 3LIP) were solved in

the open conformation (Figure 6B) [98]. In 2000, the first X-ray structure of *P. aeruginosa* lipase A in the open conformation (*Pa*LipA$_o$) in complex with octyl-phosphinic acid 1,2-bis-octylcarbamoyloxy-ethyl ester (OCP) inhibitor (PDB code: 1EX9) was reported at 2.54 Å resolution (Figure 6C) [17].

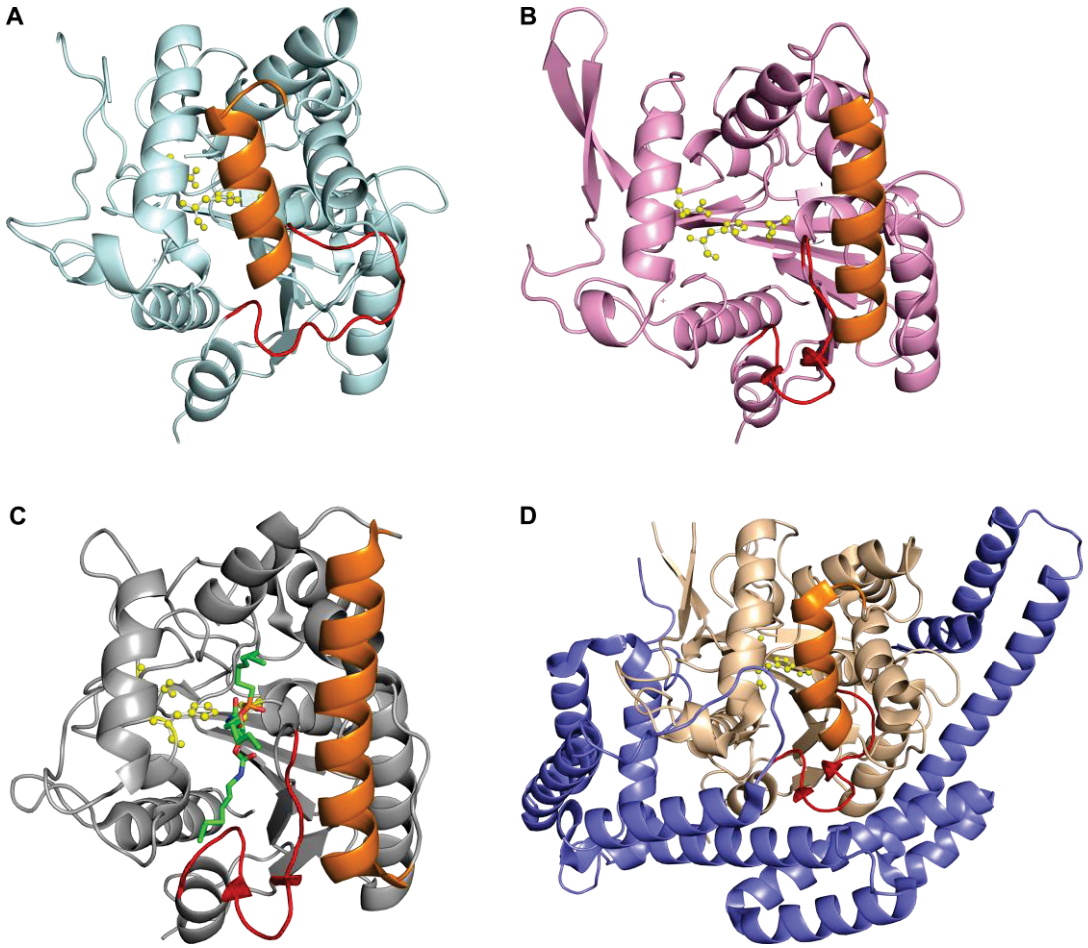


Figure 6. Crystal structures of bacterial lipases in the closed and open conformation. The catalytic triad is indicated in yellow (ball and stick representation), helix 5 (lid) is shown in orange, and residues 17-30 in red. **(A)** *Bg*LipA$_c$ (PDB code: 1QGE) [96], with helix α5 occluding the active site. **(B)** *B. cepacia* lipase A (PDB code: 3LIP) [98] crystallized in an open conformation without bound ligand, in an organic solvent. **(C)** *Pa*LipA$_o$ (PDB code: 1EX9) [17], crystallized with a bound inhibitor (OCP, shown in green sticks) in the open, active site. In both (B) and (C), helix α5 moved away from the catalytic triad residues, in order to expose the binding site, and a partial β-sheet structure is present in the region of residues 17-30 (red) to form the cleft of the binding site. **(D)** Crystal structure of *Bg*LipA (wheat) in complex with its specific foldase *Bg*Lif (blue) (PDB code: 2ES4) [19]. In the complexed lipase A, helix α5 occludes the binding site although residues 17-30 show the partial β-sheet structure.

The structure of *Pa*LipA$_o$ showed a very high structural similarity to the lipase structures of the subfamily I.2. As all lipases, it also exhibits a characteristic folding pattern of the core domain and possesses two stretches of residues, respectively ranging from 1 to 108 and 164 to 285,

which form the α/β-hydrolase fold. The active site cleft consists of four α-helices (α4, α5, α6, and α8) and a partial β-sheet structure (residue 17-30) [17, 80]. Helix α5 (residue 109-163) and its neighboring loops form a 'lid' which shields the active site in *Cv*LipA$_c$ and *Bg*LipA$_c$. In *Pa*LipA$_o$, the lid moves away from a closed to an open state in order to make the active site cavity accessible to the substrate. In the crystal structure of *Pa*LipA$_o$, an additional small antiparallel β-sheet (named as β-strand b1 and b2 in the crystal structure) is formed between β3 strand and helix α1 in the region of residues 17-30 which was absent in *Cv*LipA$_c$ and *Bg*LipA$_c$. Later in 2006, the structure of *Bg*LipA in complex with its specific foldase *Bg*Lif was crystallized in a closed conformation (PDB code: 2ES4) in which the active site was found occluded by helix α5 of the lid domain (Figure 6D) [19]. The core structure was found virtually identical to *Bg*LipA$_c$ (PDB code: 1QGE) with a root mean square deviation (RMSD) of 0.5 Å for Cα atoms [19], except for the partial β-sheet structure present in the region of residues 17-30, which is also present in *Pa*LipA$_o$.

All lipases belonging to the α/β-hydrolase superfamily have a catalytic triad formed by a nucleophilic Ser, a His and a catalytic acid-base residue that is either an Asp or Glu [11]. In the crystal structure of *Pa*LipA$_o$, the Oγ atom of Ser82 is covalently bound to the OCP inhibitor and 3.0 Å away from the Nε2 atom of the catalytic histidine H251. The carbonyl oxygen of His251and amide nitrogen atoms are hydrogen bonded to His81 and the catalytic residue Asp229. Asp 229 stabilizes the positively charged His251 during catalysis [17]. Located at a distance of approx. 15 Å from His251, a calcium (Ca) binding pocket comprises a Ca^{2+} ion coordinated in an octahedral geometry with four protein oxygen atoms and two water molecules. It is also known that all lipases possess a conserved oxyanion hole involved in stabilizing the negative charge of the tetrahedral transition state intermediate formed during catalysis [17]. In *Pa*LipA$_o$, residues Met16 and His83 belong to both the oxyanion loop and the active site of *Pa*LipA [17]. In the crystal structure of *Pa*LipA$_o$, the main chain amide groups of

Met16 and His 83 are in direct interaction with the phosphorous oxygen atom of OCP. Residues Cys235 and Cys183 form a disulfide bond [17].

The lid conformation change observed in the 3D crystal structures of lipases in closed and open conformation provides an elegant explanation of interfacial activation. Upon binding to the interface, the lid, which covers the active site, moves away from a closed to an open state, and exposes a large hydrophobic active site. However, later, it was found that lipases from *P. aeruginosa*, *B. glumae*, and *Candida antarctica* do not show interfacial activation but still have an amphiphilic lid [83]. This implies that the presence of a lid is not correlated with interfacial activation [2] but also raises the question about the nature of the activation mechanism of lipases that do not show interfacial activation.

2.4.1 *Pa*LipA biogenesis

*Pa*LipA biogenesis is a multi-step process ranging from gene expression to extracellular secretion of folded *Pa*LipA (Figure 7). *Pa*LipA gene regulation is controlled by the quorum sensing mechanism, which is a cell density-dependent regulation process [99, 100]. Extracellular signaling molecules called autoinducers are produced by many gram-negative bacteria which bind to transcriptional regulatory proteins and can induce or cease to repress specific target genes. The details for transcriptional regulation of the *Pa*LipA gene can be found in refs. [100-102]. It was found that the lipase gene LipA in *Bacillus subtilis* was encoded in an operon with a second gene LipB, which was shown to be necessary for the lipase activity as it assists the correct folding and later on was named Lif to constitute a unique class of lipase-specific chaperones [83].

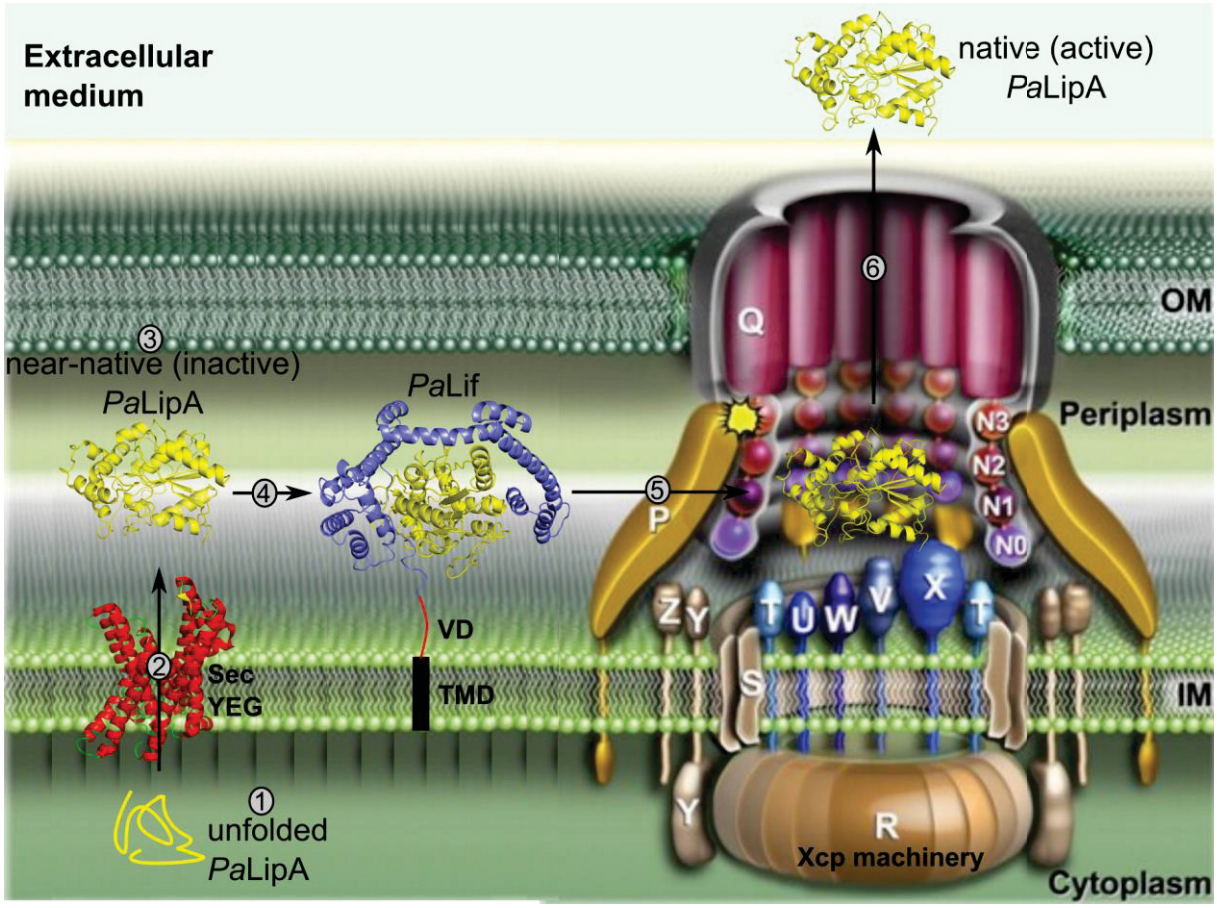


Figure 7. Proposed secretion mechanism of *Pa*LipA. (1) *Pa*LipA is synthesized in the cytoplasm with an N-terminal signal sequence that is recognized by the Sec translocon (SecYEG). (2) SecYEG transports unfolded *Pa*LipA into the periplasm. (3) In periplasm, *Pa*LipA folds to its near-native inactive conformation. (4) The inner membrane-bound *Pa*Lif binds to near-native inactive *Pa*LipA with its catalytic foldase domain (blue). (5) Upon binding with *Pa*Lif, *Pa*LipA undergoes conformational changes to convert into its active conformation. (6) Active *Pa*LipA is transported to the extracellular medium via the Xcp machinery. This figure is based on a previously published model of Xcp and lipase secretion in refs. [9, 41].

*Pa*LipA is synthesized with an N-terminal signal sequence as a precursor which targets it to the Sec translocon where SecB keeps the nascent chain in an unfolded conformation [6]. After getting exported into the periplasm, the signal sequence is cleaved and enables *Pa*LipA to fold into the secretion-competent conformation. The periplasmic maturation of *Pa*LipA is a prerequisite for its secretion to OM. In order to acquire a correct folding prior to its secretion, *Pa*LipA needs a specific intermolecular folding catalyst, which is known as 'lipase-specific foldase' *Pa*Lif [83]. In addition to *Pa*Lif, *Pa*LipA also needs 'accessory folding catalysts' like DsbA and DsbC, which catalyze the formation and isomerization of a disulfide bond [103, 104]. Other proteins are also required to enforce the correct isomeric state of the peptide bond. These

proteins include, for example, the peptidyl-prolyl cis/trans isomerases (PPIases), which catalyze the cis-isomerization of the peptide bond between Gln257 and Val258, which is a conserved feature of lipases [105]. Although *Pa*LipA variants with one or both Cys mutations were still able to fold into their active conformation, they are more prone to proteolytic degradation in the periplasm and show less efficient secretion [103, 104]. Once *Pa*LipA is folded to the native (or secretion-competent) conformation, it is transported by the Xcp-secreton across the OM. The full mechanism involving the recognition and binding of the Xcp-secreton to the native *Pa*LipA before its secretion is however still unknown.

2.4.2 Protein folding landscape

Anfinsen stated that protein folding is based on the search of the global free energy minimum conformation corresponding to the most stable 3D structure [106]. Many small proteins adopt a native conformation spontaneously that has a lower free energy than their unfolded state (Figure 8A) [107]. However, for some multidomain proteins, achieving their native and active conformation is difficult due to a high energy barrier controlling the transition from the unfolded to the native state (Figure 8B and 8C) [107]. In the case of specific bacterial proteins, such as α-lytic protease (αLP) and subtilisin, it was observed that their native conformation is comparably stable to that of their respective unfolded conformation [108, 109]. In the native protein conformational energy landscape, the folding free energy barrier for αLP is higher than 26 kcal/mol, with a folding half-time $t_{1/2} \approx 1700$ years [109]. This example demonstrates the existence of a high energy barrier which causes the folded native state to be trapped in a free energy minimum, therefore limiting the kinetics of the unfolding process while also posing folding problems [109]. The authors of this work also suggested that αLP folds spontaneously into an inactive partially folded intermediate state and requires a steric chaperone propeptide to achieve its native state [13]. It has been reported that this steric chaperone plays a crucial role

in the folding of protease by lowering the folding free energy barrier and imprinting the required steric information to the protease [13, 107, 109].

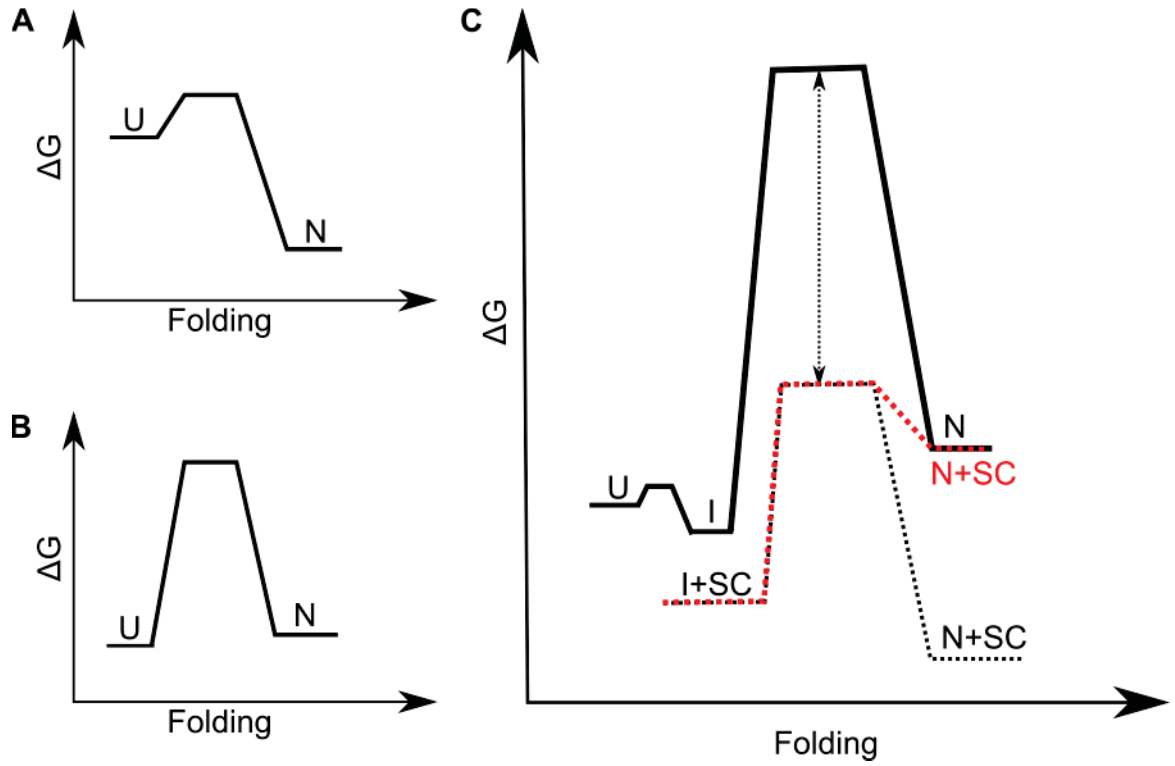


Figure 8. Protein folding landscape. (A) Small unfolded proteins (U) spontaneously fold into their native (N) low free energy conformation reversibly, without any helper proteins on relevant biological timescales. **(B)** Some multidomain or large proteins do not fold themselves into the native conformation due to the high free energy barrier between unfolded and native state. Usually, the native state of these proteins has a similar free energy as the unfolded state. **(C)** Generally, these proteins fold spontaneously to an intermediate state (I) having lower free energy compared to the unfolded state. In order to fold into the native state, this intermediate state needs a helper protein like a steric chaperone (SC), which lowers the free energy barrier (black-dotted line), thereby increasing the folding rate, as well as providing the required steric information to the protein, which increases the folding efficiency. After removal of the steric chaperone, the fully folded protein remains trapped by a high unfolding free energy barrier (black). Moreover, sometimes, steric chaperone increases the folding efficiency by lowering the free energy barrier and traps the target protein into a high energy folded state (red-dotted line). Figure adapted from ref. [107].

El Khattabi and his co-workers showed that the native state of *Bg*LipA is highly stable, as very harsh conditions were required for its denaturation [110]. They also showed that in the absence of *Bg*Lif, denatured *Bg*LipA folds to an inactive, near-native intermediate state, while upon addition of *Bg*Lif, immediate activation of intermediate *Bg*LipA was observed [85]. These observations suggested that a large energy barrier also exists along the *Pa*LipA folding pathway, which prevents the unfolding of the kinetically stable native *Pa*LipA. Therefore, it

suggests that *Pa*Lif behaves as a steric chaperone, which lowers the folding barrier by embossing the required steric conformational changes on *Pa*LipA [20]. Based on structural similarities with the secreted αLP-propeptide system, it was assumed that *Pa*LipA also follows the similar folding pathway. Nevertheless, near-UV circular dichroism (CD) and Trp-fluoresecence spectra showed that native *Bg*LipA significantly differs from that of intermediate *Bg*LipA. Investigating and understanding the complete folding and secretion mechanism of lipases is a huge challenge as it demands the in-depth characterization of the intermediates and transition folding states, but also requires to determine the kinetics of their formation in the folding process, as well as the nature of their respective interactions with Lif and the Xcp machinery. **In publication III**, we have provide insights into *Pa*LipA folding from the binding intermediate to the native state in the presence and absence of *Pa*Lif.

2.5 *Pa*Lif and its importance in *Pa*LipA biogenesis

Other than a general folding catalyst, the periplasmic folding of *Pa*LipA critically depends on the lipase-specific foldase, *Pa*Lif. Usually, foldases are encoded in the same operon together with their cognate lipases (**section 2.4**) [100], and aside from their genome sequences, not much experimental data is available about their function. Lifs are classified into four families (I-IV). Lifs from *P. aeruginosa*, *Pseudomonas mendocina*, *Pseudomonas wisconsinensis,* and *Pseudomonas alcaligenes* belong to family I, Lifs from *B. cepacia*, *B. glumae*, *Pseudomonas fragi*, *Xylella fastidiosa,* and *Ralstonia metallidurans* represent family II, family III Lifs are from *Acinetobacter calcoaceticus*, and family IV contains Lifs from *Pseudomonas* sp. strain *KFCC10818*, *Vibrio cholerae* and *Vibrio vulnificus* [9].

Regardless of their low sequence similarity, all Lifs share a very similar secondary structure composed of 70 % α-helical and 30 % random coil elements, which indicates that the catalytic mechanism of Lifs is also conserved [9]. In 2006, the soluble form of *Bg*Lif in complex with its cognate lipase was crystallized (PDB code: 2ES4) (Figure 9) [19]. All Lifs contain an N-

terminal hydrophobic segment that anchors them into the bacterial IM. The remaining part of the foldase is exposed to the periplasm. Upon truncation of the N-terminus, *Pa*Lif was still able to catalyze the folding of its cognate lipase, indicating that the membrane anchor is not needed for the folding activity of *Pa*Lif *in vitro* [111]. It was suggested that the membrane anchor of *Pa*Lif is required to prevent its secretion in complex with *Pa*LipA by the Xcp machinery [80]. In *Pa*Lif, the membrane anchor is followed by a stretch of proline- and alanine-rich segment known as variable domain. This region accounts for about 20 % of the entire *Pa*Lif sequence and is structurally disordered. This region presumably acts as a spacer that establishes a certain distance between the C-terminal domain and the IM, so that the C-terminal domain can sufficiently protrude into the periplasmic space [9, 80, 112]. The variable domain is followed by a large C-terminal domain known as catalytic folding domain. Many experiments suggest that solely the C-terminal domain interacts with the lipase A and assists its folding [80]. The crystal structure of *Bg*Lif:LipA revealed the previously unobserved fold of the *Bg*Lif catalytic domain. It consists of 11 α-helices forming an open "head-phone" like structure, which engulfs *Bg*LipA. These 11 α-helices are organized into three domains: two globular mini-domains (MD1 and MD2) connected by an extended helical domain (EHD). MD1 is composed of three N-terminal helices (α1-α3) of the catalytic domain, and MD2 is composed of three C-terminal helices (α9-α11). They are located on opposite sides of *Bg*LipA connected by the EHD which contains four α-helices (α4-α8). Importantly, all aromatic residues are located within these mini-domains [19]. Out of 11 α-helices, six are in direct contact with *Bg*LipA, thereby forming a large interface (~ 5400 $Å^2$ buried solvent accessible surface area (SASA) of the interface) [19]. It was found that a stretch of eight amino acids, RxxFDY(F/C)L(S/T)A (where x represents any amino acid), is highly conserved in all Lifs [9]. This conserved sequence motif was assigned to the helix α1 of *Bg*Lif that directly interacts with *Bg*LipA in the crystal structure (Figure 9).

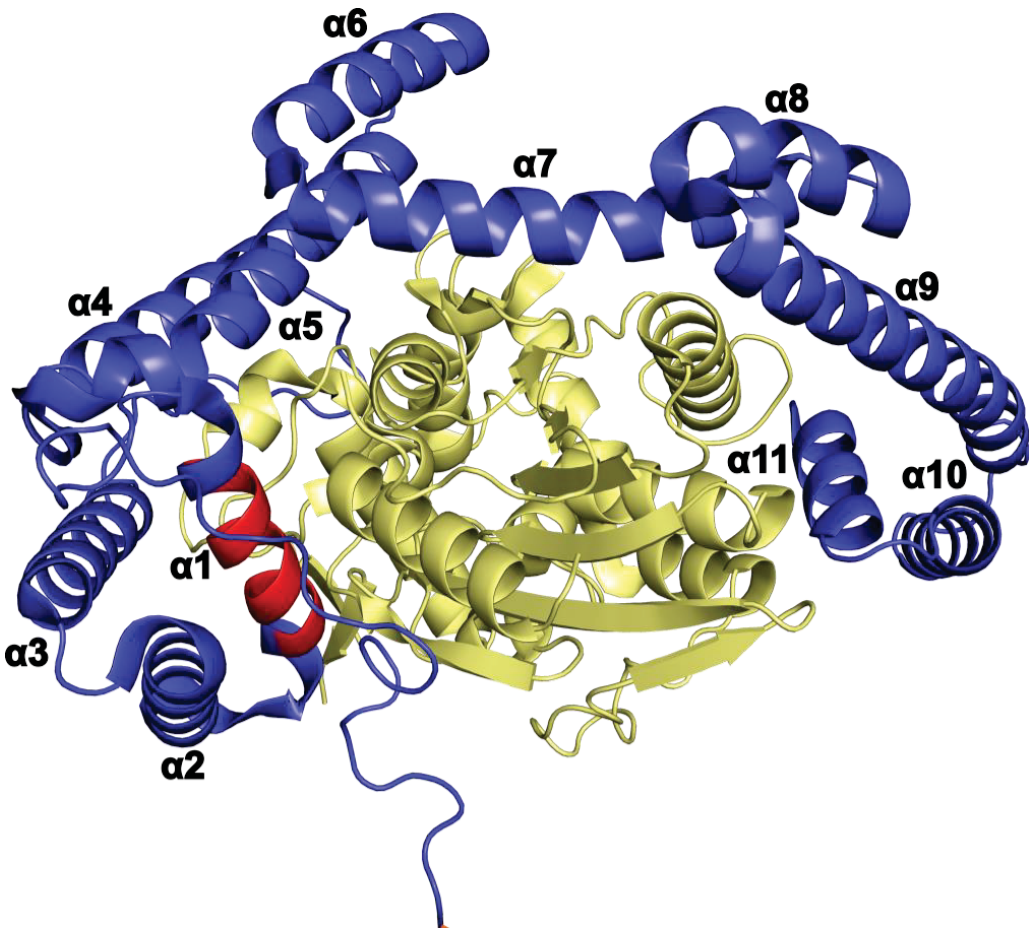


Figure 9. Crystal structure of the catalytic folding domain of *Bg*Lif in complex with its cognate lipase. The structure of the catalytic folding domain (blue) of *Bg*Lif forms a unique "head-phone like" fold to engulf its cognate lipase *Bg*LipA (yellow) (PDB code: 2ES4) [19]. The C-terminal catalytic folding domain of *Bg*Lif contains 11 α-helices of which six are in direct contact with *Bg*LipA. The conserved motif present on helix α1 is shown in red.

Mutation of amino acids Y99 and S102 located in the conserved foldase motif of *Pa*Lif resulted in foldase variants that were unable to activate *Pa*LipA [113]. Mutation Y99 is further studied in **publication II**, to understand the possible role of this residue in *Pa*Lif-induced *Pa*LipA activation. Moreover, eight intermolecular interactions were identified at the *Bg*Lif:LipA interface between evolutionarily conserved amino acids of the Lif family (E107, R214, Q253, Q264, R288, R343, S346, and R349) and the lipase family. All of the eight amino acids mentioned above are located on the catalytic domain of *Bg*Lif. Pauwels *et al.* reported that both *Bg*Lif and *Bg*LipA undergo structural changes upon binding, mainly in the EHD of *Bg*Lif [19]. They also proposed that R343 of *Bg*Lif interacts with E63 of *Bg*LipA, thereby contributing to the specificity of *Bg*Lif to bind its cognate lipase [19]. Several *in vitro* experiments have shown

that *Pa*Lif forms a very stable complex with its cognate *Pa*LipA with K_D = 5 nM [85, 111]. Usually, a lipase gene and its cognate foldase form an operon suggesting a 1:1 ratio for both lipase and foldase expression, but in *P. aeruginosa* and *B. glumae*, foldase was produced in significantly lower amounts than their cognate lipase [85]. Lifs have been known to specifically activate only their cognate lipases as no lipase folding occurred when *Pa*Lif is coexpressed with *Bg*LipA [114]. However, a complete understanding of the physiological role of foldases is still unknown.

2.5.1 Molecular chaperones

Proper folding and localization of proteins are of extraordinary importance within each cell [115]. The deposition of misfolded proteins into amyloids can cause neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, and Huntington's disease [115], which signifies the importance of proper protein folding. Protein aggregation is disastrous for cells as it triggers self-induced cell death, also known as apoptosis. Many studies indicate that incorrect folding is one of the main reasons for protein aggregation [116]. In order to prevent or reverse the protein aggregation caused by misfolding, cells use molecular chaperones as protein folding assistants.

Usually, a polypeptide tries to adopt a distinct three-dimensional (3D) structure under normal physiological conditions. A multitude of interactions are involved in this folding process [117]. Nonetheless, the same molecular interactions can also result in misfolding of proteins. Therefore, several helper proteins are present in the cell to prevent the formation of wrong interactions, known as classical chaperones. Initially, they were identified as stress proteins or heat shock proteins (HSPs) that are expressed upon a transient increase of temperature [118]. HSPs are categorized into two groups, namely 'protein holding' and 'protein folding' [119]. HSPs, which belong to the protein holding category, bind to the partially unfolded protein during heat shock and prevent the exposure of hydrophobic surface. Chaperones which catalyze

protein folding usually self-associate to form large folding chambers in which the substrate protein can undergo the required conformational changes by forming intramolecular interactions [120]. Therefore these classical chaperones increase the folding efficiency for their target proteins but do not affect the folding rate constant [121].

However, some proteins fold to an intermediate, non-native state and fail to fold into their native state on a biologically relevant timescale because of a huge free energy barrier between non-native and native state (**section 2.4.2**) (Figure 8B). These proteins usually need the assistance of highly specific steric chaperones that imprint the required conformational information onto the trapped intermediate (non-native) state while lowering the folding energy barrier (Figure 8C) [107]. Until now, very little information is available about these highly specific steric chaperones. In addition to Lif, few other bacterial steric chaperones have been identified over the last few years and studied in detail, such as propeptides (also called prodomains) which are required for the folding of proteases and fimbrial chaperones required for the folding of subunits of adhesive pili (Figure 10) [107].

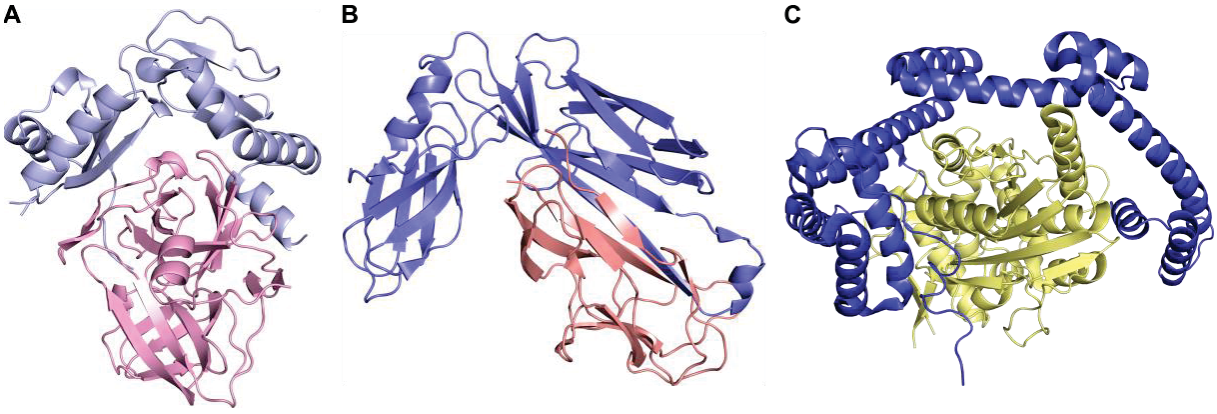


Figure 10. Crystal structures of steric chaperones in complex with their target proteins. (A) Crystal structure of αLP (pink) in complex with its propeptide (light blue) from *Lysobacter enzymogenes* (PDB code: 4PRO.pdb) [122]. **(B)** Crystal structure of fimbrial chaperone (medium blue) from *Escherichia coli* interacting with the FimH pilin domain (salmon) (PDB code: 1ZE3) [123]. **(C)** *Bg*Lif (dark blue) in complex with its cognate lipase (yellow) (PDB code: 2ES4) [19].

Bacterial extracellular proteases and propeptides are synthesized together with the signal peptide in the cytoplasm. Upon translocation through the IM, the signal peptide is cleaved off

whereas the propeptide is retained with the protease. The propeptide prevents premature activation of the protease, thereby promoting the correct folding [124]. Well-studied examples of such bacterial proteases are αLP and subtilisin produced by *B. subtilis* (**section 2.4.2**). It has been observed that in the absence of propeptide, subtilisin and αLP can fold into molten globule intermediates that are unable to fold further into the enzymatically active conformation [125]. Addition of propeptide can catalyze the conversion of this molten globule intermediate into the active native state. After activation, the propeptide is removed by autoproteolysis, in order to liberate the active protease which is kinetically trapped in the folded native state due to high unfolding energy barrier.

Adhesive pili in gram-negative bacteria require the periplasmic chaperone PapD for the assembly of adhesive filaments on the bacterial surface. Subunits of adhesive pili have an incomplete immunoglobin-like fold, which creates a deep hydrophobic groove on their surface due to a missing last β-strand [126]. The periplasmic fimbrial chaperone binds to the non-native subunits and provides the missing β-strand while catalyzing the folding of the subunit [126, 127]. In the absence of the chaperone, the non-native subunits aggregate and are degraded by the DegP protease [128]. Unlike the propeptide, after folding to the native state, fimbrial chaperones remain bound to the native subunits to direct them to the OM assembly platform [127]. In contrast to other steric chaperones, the fimbrial chaperones do not fold their substrate into their final, native fold but instead trap the subunit in a high energy folding intermediate.

Therefore, regardless of structural similarities, it appears implausible that these steric chaperones have a common mechanism of action.

2.6 Methods

2.6.1 Homology modeling

The generation of the 3D structure of a protein (target) based on the known experimental structure of homologous proteins is referred to as homology modeling [129-131]. Usually, proteins from the same family show high conservation of the structure compared to the conservation of the sequence. Therefore, we can use homology modeling to create relatively accurate protein models using template structures [129-132]. Usually, the first step in homology modeling is to identify templates and aligning the sequence of the target to each template according to the similarity of their residues known as "multiple sequence alignment". A Correct alignment is a crucial step as it gives an overview of the features of the protein family such as conserved motifs, insertions, and deletions. An incorrect alignment leads to errors in the predicted structure. After the correct sequence alignment, the model building part begins, in which the backbone of the target protein is built in the same position of the template residues by using the alignment. This is followed by building missing parts such as loops, adding side chains, optimization of conformations of side chains and finally, energy minimization of the built model.

As no crystal structure is available for closed *Pa*LipA and free *Pa*Lif, homology models were built for the same. All lipases from family I, which have been crystallized so far in the closed conformation, exhibit a very similar fold (**section 2.4**), with sequence identities of > 41% and sequence similarities of > 73%, which make it possible to model *Pa*LipA in a closed conformation. By contrast, for Lif, only a single crystal structure is available, which is from *B. glumae* in complex with its cognate lipase. It has a sequence identity of 39% and 52% sequence similarity to *Pa*Lif. However, all Lifs possess a conserved motif of eight amino acid in MD1 (**section 2.5**) that can help in building the correct model. Furthermore, at the lipase:foldase interface from *B. glumae*, eight intermolecular interactions were identified between amino

acids, which are evolutionarily conserved in the foldase family and also in the lipase family that guide the modeling of the complex.

2.6.2 All-atom MD simulations

One of the principal and state-of-the-art methods to study the dynamics of biological molecules is the computational method of MD simulations. Studying protein dynamics *in vivo* is challenging, expensive, and time-consuming, because of the complex dynamic structure of proteins [133]. In MD simulations, the time-dependent behavior of a molecular system based on Newton's laws of motion is calculated [134]. MD simulations are now consistently used to explore the dynamics and thermodynamics of the system of interest, as well as in the determination of X-ray crystallography and NMR structures [135] (**publication II)**. MD simulations generate information at the atomistic level, including atom positions and velocities and helping to connect microscopic lengths and time scales with the macroscopic world of experiments.

All-atom MD simulations have evolved among the different computational methods, being the tool of choice to analyze the conformational dynamics and fluctuations of a molecule over time [136-138]. As one of the main objectives here is to understand the conformational dynamics of *Pa*Lif and *Pa*LipA at the atomic level, we used all-atom MD simulations throughout this research. Despite the utility of MD simulations, they are limited by two factors: the force field used and the high computational costs. This often results in an inadequate sampling of conformational space, especially in the case of intrinsically disordered proteins (IDPs) [139, 140]. The high computational cost has been partially tackled with the nowadays available computational resources and technological developments (i.e., GPU-enabled MD code and hardware development), making it possible to run long timescale simulations up to the microsecond scale [141]. However, the time scale in which free *Pa*Lif undergoes reversible changes between compact and extended states of the head-phone like structure

(~ 100 microseconds) (**publication I**), is still beyond the timescale reached by all-atom MD simulations. Keeping long timescale aside, we ran multiple independent all-atom MD simulations of 1 μs length, which provide the overall trend of conformational changes for the free *Pa*Lif, *Pa*LipA (open and closed) and *Pa*Lif:LipA complex [142, 143].

2.6.3 Maximum entropy refinement method

In order to combine experiments and simulations, the maximum entropy method (MEM) is the plausible choice. As mentioned above (**section 2.6.2**), force fields have limited accuracy, and therefore, MD simulations of macromolecules sometimes produce results that are not complete and quantitatively in accordance with experiments. MEM is a method to obtain the most probable distribution that best represents the state of knowledge, given a set of experimental data [144, 145]. Among the infinite number of distributions that are compatible with the data, one should select the distribution which maintains the largest degree of uncertainty about the variables of interest [144]. Shannon's entropy is the natural quantity for expressing uncertainty in a given distribution,

$$S(p) = -\sum_i^n p(x_i) \ln p(x_i) \qquad \text{(eq. 1)}$$

where p is the probability of some testable information i about a discrete quantity x [146]. In order to find the best probability distribution which is compatible with the observed data, this expression should be maximized under the constraint that $p(x)$ sums to 1. This principle has been extended to the maximum relative entropy (MRE) principle, eq. 2 [145, 147]. The advantage of MRE principle is being invariant with respect to posterior distribution such that, it plays an important role in multiscale problems. The entropy is here computed relative to a given prior distribution $p_o(x_i)$.

$$S\,[p|p_o] = -\sum_i^N p(x_i) \ln \frac{p(x_i)}{p_o(x_i)} \qquad \text{(eq. 2)}$$

By maximizing this expression under the experimental constraints, we find the posterior distribution that is as close as possible to the original distribution p_o, but now more compatible with the experimental data. The problem of maximum entropy in the context of noisy data has been addressed numerous times in other fields [148], but the approach given by Gull and Daniell is potentially useful in MD simulations [149]. In the case of multiple constraints, they can be replaced with a single constraint χ^2 on the statistic over all data, only matching them up to their experimental uncertainty ($\chi^2 \leq 1$), given by eq. 3

$$\chi^2 = \frac{1}{M}\sum_j^M \frac{(D_j^{sim} - D_j^{exp})}{\sigma_j^2} \qquad \text{(eq. 3)}$$

Here, M is the number of data points, D_j^{exp} is the set of experimental measurement and σ_j^2 is the experimental error for the parameter j. D_j^{sim} is the corresponding quantity calculated from MD simulations. To quantify the change in the posterior distribution under the experimental constraints and in maximum entropy, it can be associated with a free energy function [21, 150].

*Pa*Lif is minimalistically structured and an underdetermined system with respect to conformational changes in free form. Therefore, to gain insights at the atomistic level as to the conformational variability of free *Pa*Lif, we reweighted the MD simulations-generated ensemble. I applied the MRE principle under experimental χ^2 constraints in **publication I**. Experimental constraints are obtained from fluorescence spectroscopy measurements preformed in the lab of Prof. Dr. C. Seidel.

2.6.4 Förster resonance energy transfer

FRET can be used to measure the distance between two fluorophores, which provides structural and dynamical insights into the conformational landscape of proteins with high spatial and temporal resolution [151]. FRET is a mechanism of energy transfer from a donor fluorophore to an acceptor fluorophore through non-radiative dipole-dipole coupling [152, 153]. The donor

fluorophore, initially in its electronically excited state, emits energy while returning to the ground state [153]. This emission spectrum corresponds to the excitation spectrum of an acceptor fluorophore [153]. The rate of energy transfer from donor to acceptor is inversely proportional to the sixth power of the distance between donor and acceptor fluorophore (eq. 4). Therefore, FRET is a strongly distance dependent non-radiative energy transfer method. [154].

$$k_{\mathrm{FRET}} = \frac{Ro^6}{\tau_{D(0)}\, R_{\mathrm{DA}}^6} \qquad \text{(eq. 4)}$$

Where $\tau_{D(0)}$ is the donor lifetime, R_{DA} is the distance between donor and acceptor fluorophore and *Ro* is the Förster radius where the transfer efficiency is 50% [153]. FRET can also be quantified by FRET efficiency, which is the number of quanta transferred from donor to acceptor divided by total number of quanta absorbed by donor. The distance R_{DA} can be calculated from the donor lifetime and the FRET efficiency [152, 153, 155, 156]. The FRET rate constant also depends on the orientation of the dyes and distance, and it is most sensitive in a distance range of 20 Å to 150 Å [157]. This approach has frequently been used to determine donor-acceptor distance distributions $p(R_{\mathrm{DA}})$ within one structure or between multiple structures to understand the dynamical behavior in solution [157-159]. FRET spectroscopy can accurately track distances over timescales ranging from ns to seconds [160]. Therefore, FRET was the method of choice to study the conformational dynamics of free *Pa*Lif in solution in **publication I**. The distance distribution profile and χ^2constraints obtained by FRET (**section 2.6.3**), were used for reweighting of the large structural ensemble of free *Pa*Lif generated by MD simulations in **publication I**. In order to have corresponding data from MD simulations (**as mentioned in section 2.6.3**), the accessible volume (AV) of the dyes was calculated for all the conformations generated from MD simulations. AV considers the dyes as hard sphere models connected to the protein via a flexible linker and represents uniform population densities of the spatial dye position [151, 155, 156]. These densities can be used to calculate FRET observables, in this case, the distance distributions for a large ensemble of structures generated

by MD simulations. This method is successfully applied in **publication I** to obtain corresponding distance distributions for the ensemble of free *Pa*Lif conformations generated from MD simulations. Further details about AV calculations of dyes can be found in refs. [151, 155, 156].

Depending on the complexity of the molecule or system, FRET experiments can be performed for single molecules, sub-ensembles (selectively averaged single-molecule events), and ensembles [157]. Single-molecule FRET has the advantage that it allows one to resolve distributions of FRET observables and to obtain kinetic information at the same time [157, 161, 162]. Hence, multiple distinguishable static states and dynamics of the states can be obtained [163].

2.6.5 Potential of mean force computation using umbrella sampling

The free energy difference is the driving force of any process under physiological conditions, such as the folding of a protein [164]. Umbrella sampling (US), i.e. biased MD simulations, is one of the efficient methods that provides a free energy profile of a process along a reaction coordinate (ξ) (Figure 11) [165-167]; this is the method used in **publication III**. The reaction coordinate may be composed of a single or multiple parameters such as distance, angle, or RMSD [167, 168]. In US, an additional energy term or bias is applied to the system to ensure efficient sampling along the reaction coordinate. During US, the pathway along the reaction coordinate is divided into a series of windows with a defined reference point (ξ^{ref}) [167, 168]. The bias potential w_i of the window i is often a simple harmonic bias of strength K (eq. 5) [167]. This ensures effective sampling in all regions, split into several windows, and overlapping distributions.

$$w_i(\xi) = K/2(\xi - \xi_i^{ref})^2 \qquad \text{(eq. 5)}$$

The free energy profile can be obtained as a potential of mean force (PMF) [166, 168], which defines the free energy F as a function of a reaction coordinate ξ (eq. 6)

$$F(\xi) = -k_B T \ln[P(\xi)] + C \qquad \text{(eq. 6)}$$

where k_B is the Boltzmann factor, T is the temperature, C is the constant and $P(\xi)$ is the probability of the system along the reaction coordinate [169, 170].

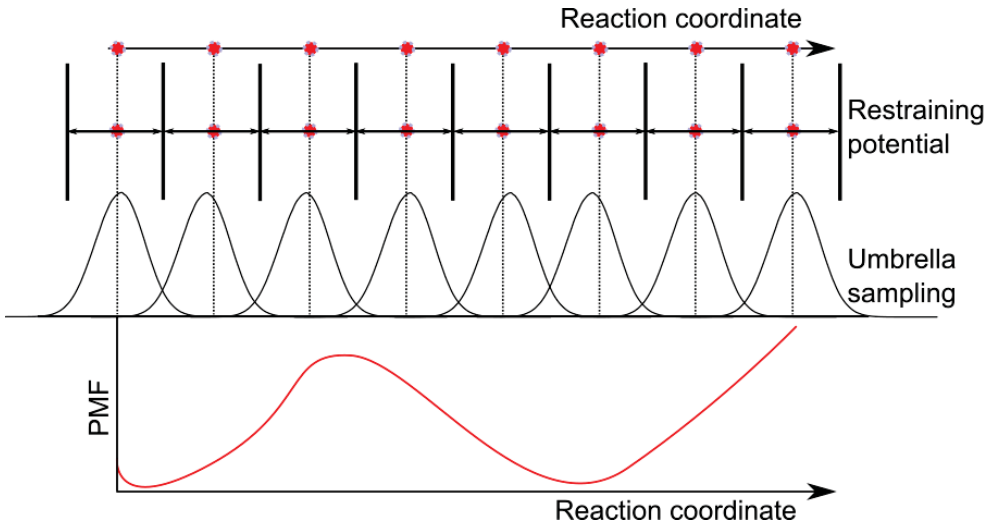


Figure 11. Schematic representation of umbrella sampling method for PMF computations. First, the reference configurations (red dots) are generated along the reaction coordinate. Then these configurations are restrained by harmonic potential during MD simulations which result in overlapping umbrella sampling distributions. Finally, the free energy profile (PMF) can be obtained using WHAM.

Next, to estimate the unbiased, free energy profile for combined US along the reaction coordinate, many methods have been proposed, but the most reliable approach is the Weighted Histogram Analysis Method (WHAM) [171]. Details about WHAM can be found in ref. [172]. In brief, in WHAM, first the unbiased distribution for each window *i* is obtained according to eq. 7

$$P(\xi)_i^{unbiased} = P(\xi)_i^{biased} \exp\left(-\frac{w_i(\xi)}{k_B T}\right) \left\langle \exp\left(\frac{w_i(\xi)}{k_B T}\right) \right\rangle \qquad \text{(eq. 7)}$$

Where $\langle\ \rangle$ indicates the average over the ensemble. Following, $P(\xi)_i^{unbiased}$ are combined to calculate the global distribution by a weighted average of the distribution of the individual windows using eq. 8,

$$P(\xi)^{unbiased} = \sum_i^{windows} P(\xi)_i^{unbiased}\ p(\xi)_i \quad \text{(eq. 8)}$$

where p_i are the weights, selected in order to minimize the statistical error of $P^{unbiased}$, under the condition $\sum p_i\,(\xi) = 1$, and calculated from eq. 9

$$p_i(\xi) = N_i \exp\left(-\frac{w_i(\xi) + F_i}{k_B T}\right) \quad \text{(eq. 9)}$$

where N_i is the total number of steps sampled during window *i* and F_i is the free energy constant calculated by eq. 10

$$\exp\left(-\frac{F_i}{k_B T}\right) = \int P(\xi)^{unbiased} \exp\left(-\frac{w_i(\xi)}{k_B T}\right)\, d\xi \quad \text{(eq. 10)}$$

Eq. 8-10 have to be iterated self-consistently until convergence is reached.

2.6.6 Constraint Network Analysis

Biomolecules are generally composed by a hierarchy of stability reflecting their structural modularity. Getting insights into this hierarchy allows for understanding the relationship between biomolecular structure, (thermo) stability, and function. To this end, the Constraint Network Analysis (CNA) approach has been first introduced by S. Radestock and H. Gohlke [173, 174] and further developed into the efficient CNA software package [175]. CNA functions as a front- and back-end to the graph-theory based FIRST (Floppy Inclusion and Rigidity Substructure Topology) software [176]. Here, a biomolecule is represented as a constraint network with bodies representing atoms connected by a set of bars representing covalent and non-covalent interactions [177, 178]. Once the constraint network is built, the fast-combinatorial pebble game algorithm [179] decomposes the network into its flexible and rigid regions by counting the number and spatial distribution of internal degrees of freedom (floppy modes).

CNA allows for simulating a bond dilutions process of biomolecules by gradually removing non-covalent constraints from initial network representations [174, 177, 180, 181] (Figure 12).

To this end, the strength of each polar interaction E_{HB} is determined from an empirical energy function [182]. For a given network state $\sigma = f(E_{cut})$, hydrogen bonds (including salt bridges) with an energy $E_{HB} > E_{cut}$ (σ) are removed from the constraint network [183]. Accordingly, stronger hydrogen bonds will break later during the bond dilution process than weaker ones. Finally, rigidity analysis are performed on each constraint network state σ resulting in a bond dilution trajectory. A bond dilution trajectory is then analyzed by CNA, which calculates several global and local indices [176] Recently, an efficient ensemble-based perturbation approach has been implemented in CNA that provides an excellent tool for studying the long-range effect of altered structural stability due to amino acid substitution and ligand binding [183]. In this approach, perturbations are introduced *in silico* by removing all constraints from the network associated with the presence of a bound ligand or alanine substitution (Figure 12). Because this perturbation is introduced at the network level it does not change the biomolecular structure.

To quantify the change in structural stability upon perturbation neighbor stability maps are used (Figure 12). Neighbor stability maps record the state during the bond dilution trajectory when a rigid contact between pairs of neighboring residues cease to exists. Taken together, the local stabilities of the residue-residue contacts display the hierarchy of biomolecular stability. Recently, it has been shown that the sum over all rigid contacts represents a chemical potential energy E_{CNA} [184], due to non-covalent bonding. The difference in the chemical potential energy, $\Delta E_{CNA} = E_{CNA,perturbed} - E_{CNA,ground}$, reflects the change in biomolecular stability upon removal of the ligand or amino acid substitution. Because the introduced perturbations in the ground network state are localized and small, a one-step free energy perturbation approach can be used to compute the free energy of altered stability ΔG_{CNA} (Figure 12) [183]. A per-residue decomposition $\Delta G_{i,CNA}$ schema allows measuring the extent of each residue i to the change in

structural stability. This per-residue free energy has been successfully applied for studying pathways for signal transmission in biomolecules [183, 185].

I used this ensemble-based perturbation approach in **publication II and III** to understand how the mutation of certain residues of *Pa*Lif modulates the mechanical stability of *Pa*LipA in the complex. From these findings I intended to identify the underlying mechanism of *Pa*LipA activation mediated by *Pa*Lif.

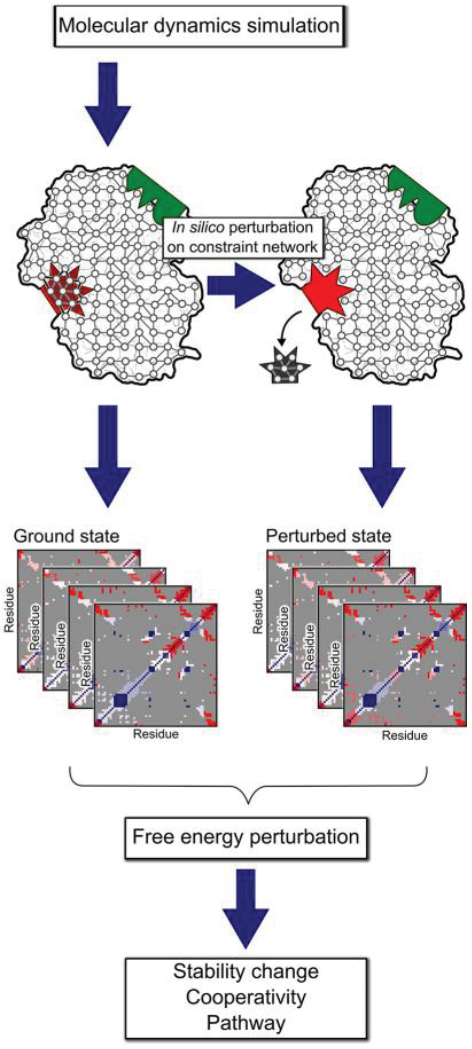


Figure 12. Schematic representation of the workflow of the perturbation approach. As input for the perturbation approach, a structural ensemble extracted from MD trajectories of the ground state is used. The perturbed ensemble is then obtained by removing all constraints associated with a bound ligand or amino acid substitution. From differences between neighbor stability maps of the ground and perturbed state, a free energy of altered stability is derived. This free energy value provides valuable insights into structural stability, allosteric cooperativity, and signal transmission in biomolecules. This figure is taken from ref. [183]

3. SCOPE OF THE THESIS

Lipases are widely used in many industries, such as the production of biopolymers and biodiesel, dye, detergent, dairy, and pharmaceutical industry, agrochemicals, cosmetics, flavors, and many more [1, 2, 83, 94] (**Section 2.3.1**). The use of lipases is quickly increasing day by day for a variety of biotechnological applications, which demands the production of lipases at a ton scale. To achieve such yields, the efficient overexpression of the corresponding genes, folding, and secretion are required [1, 2, 5, 186]. Despite that lipase A from *P. aeruginosa*, being the most well-studied "true" lipase, its heterologous production in well-studied model organisms regularly fails due to its complex secretion mechanism [1]. Furthermore, the mechanism of its proper folding is unknown to date. The most important protein in the mechanism of folding and secretion of *Pa*LipA is the lipase-specific foldase *Pa*Lif [1]. *Pa*Lif is a steric chaperone and is required to convert inactive near-native lipase A to active lipase A (**section 2.4.1**). At present, minimal information is available about the structural dynamics of *Pa*Lif in its free state, which can provide new perceptions on the binding and activation process of *Pa*LipA. Furthermore, the mechanism of the assisted folding of *Pa*LipA in the presence of *Pa*Lif is unknown. When *Pa*LipA is properly folded, whether *Pa*Lif is also involved in its secretion is also unknown. Here, the complete role of *Pa*Lif is not fully understood. This knowledge could be utilized to improve the folding and secretion of *Pa*LipA and, thus, improve its yield to satisfy the high demand in various industries. This poses the following questions:

- What is the structural and functional dynamics of *Pa*Lif in the absence of *Pa*LipA?
- What is the role of the structurally stable domain MD1 of *Pa*Lif in the *Pa*LipA activation?
- How does *Pa*Lif facilitate *Pa*LipA folding from inactive to active precursor?

Those questions have been addressed in this thesis, which led to the following publications.

4. PUBLICATION I - FUNCTIONAL DYNAMICS OF A STRUCTURALLY MINIMALISTIC CHAPERONE

Jakub Kubiak[#], Neha Verma[#], Peter Dollinger, Holger Gohlke, Claus Seidel, Filip Kovacic, Karl-Erich Jaeger

[#] both authors contributed equally to this work.

Original publication, see pages 100-161

4.1 Background

Steric chaperones directly catalyze the folding of the target protein by providing the essential steric information in the adenosine triphosphate (ATP)-independent manner in contrast to classical molecular chaperones (**section 2.5**) [107]. *Pa*Lif catalyzes the folding of *Pa*LipA from the near-native, inactive state to the open, active state (**section 2.4**) [80]. However, how *Pa*Lif assists the *Pa*LipA folding remains elusive, and a detailed description of the structural preferences of *Pa*Lif is prerequisite for understanding its mode of action. The catalytic folding domain of *Pa*Lif has a three domain organization (Figure 13A) (**section 2.5**).

In this publication, we investigated the conformational dynamics of *Pa*Lif in the lipase free state by combining high precision FRET and MD simulations. Our results shed light on the structural stability and functionality of the individual domains of *Pa*Lif, which provide valuable insights into the general mode of action of these steric chaperones.

4.2 Results

For the following sections (**section 4.2.1-4.2.4**), the text and figures have been taken from **publication I**. In this study, I performed MD simulations, analyzed the data and combined the MD simulations with FRET experiments. The results of my work are summarized in the sections **4.2.1**, **4.2.3** and **4.2.4** of this publication.

4.2.1 Structural dynamics of the individual domains of free *Pa*Lif

Based on previous CD spectroscopy experiments and secondary structure analyzes of the *Bg*Lif:LipA complexes [80], it was postulated that binding of *Bg*LipA leads to structural changes predominantly in the EHD of *Bg*Lif, while the mini-domains, MD1, and MD2 noticeably retain their structure [19, 80]. Therefore, here first, I determined the structural stability of the isolated, free MD1 and MD2 and the spacer region (EHD) of *Pa*Lif using all-atom MD simulations (Figure 13A and 13B). In the course of the MD simulations, MD1 retained almost completely the α-helical secondary structure with average α-helical content (AHC) ~72 ± 1.0% (mean ± SEM) and ~84 ± 3.0% for ff99SB and ff14SB force fields, respectively. This finding agreed well with the secondary structure content of the NMR structure of free MD1 solved in **publication II**. Both EHD and MD2 show a less persistent α-helical secondary structure in the course of the simulations (Figure 13B), as reflected in the AHC of these domains of 45.0 ± 5% (76.4 ± 1%) and 55.9 ± 9% (70.4 ± 7%) for ff99SB (ff14SB) (Table 1).

Table 1. Averaged AHC of free isolated domains of *Pa*Lif.[a]

	AHC[a] [%]	
α-helix	**ff99SB**	**ff14SB**
H 1	73.9 ± 2	84.0 ± 3
H2	75.2 ± 2	84.2 ± 3
H3	66.3 ± 0	83.6 ± 1
MD1	71.8 ± 1	84.0 ± 3
H4	75.5 ± 3	92.9 ± 1
H5	50.3 ± 0	66.1 ± 2
H6	23.2 ± 0	53.7 ± 2
H7	50.0 ± 2	85.2 ± 1
H8	26.4 ± 3	84.1 ± 4
EHD	45.0 ± 5	76.4 ± 1
H9	66.3 ± 4	77.7 ± 2
H10	76.3 ± 3	98.9 ± 1
H11	25.1 ± 1	34.8 ± 1
MD2	55.9 ± 9	70.4 ± 7

[a] The AHC per-residue calculated over three independent MD simulations using two different force fields ff99SB and ff14SB, respectively, are given as mean ± SEM. The percentage of AHC is given for individual helices as well as for domains as mean ± SEM.

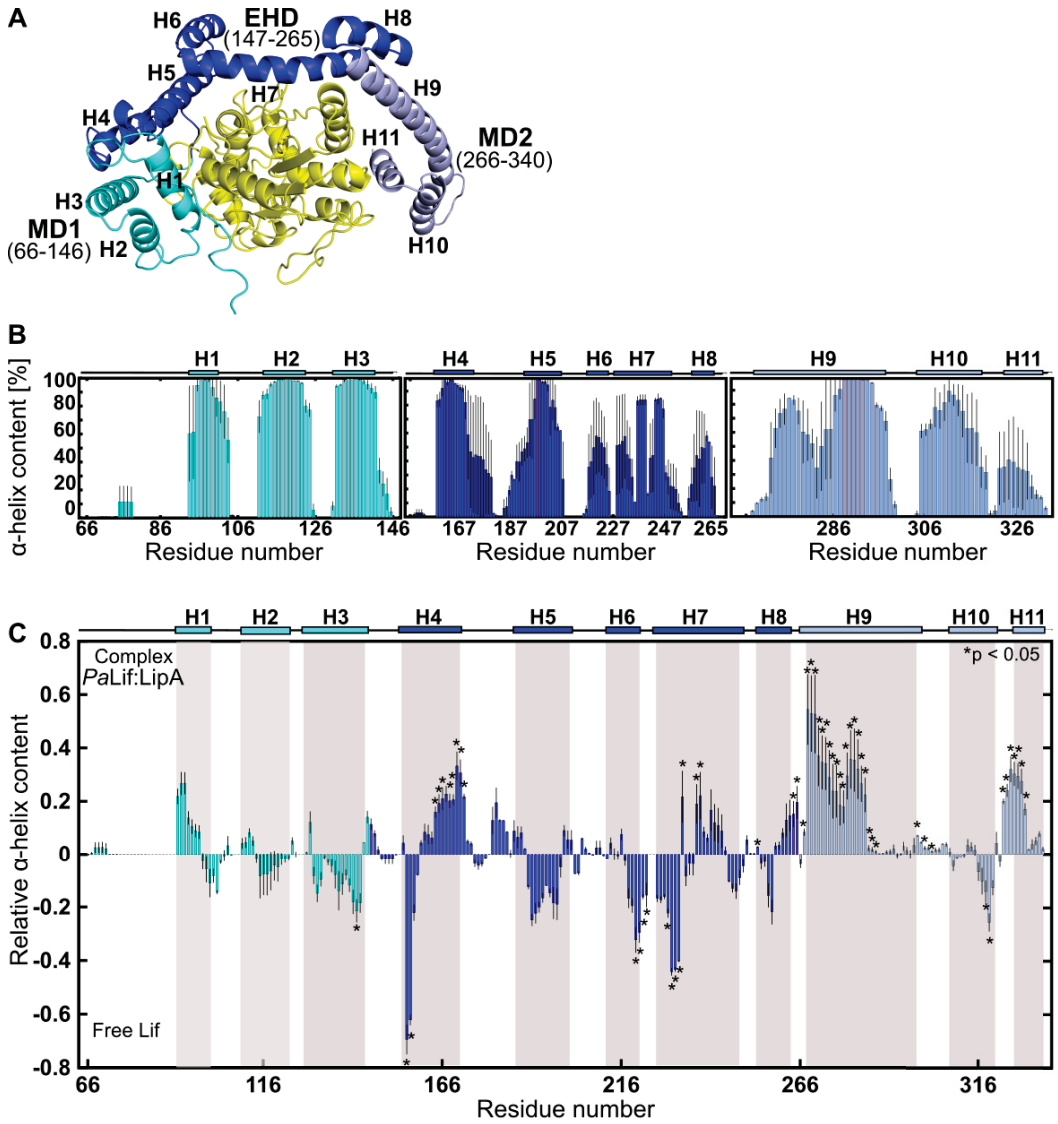


Figure 13. Secondary structure analysis of *Pa*Lif by MD simulations. (A) The tertiary structure of *Pa*Lif:LipA complex homology model showing how MD1 (residues 66-146) (cyan), MD2 (residues 147-265) (light blue), and EHD (residues 266-340) (blue) are defined. The α-helices of *Pa*Lif are labeled H1 to H11, forming the novel fold around *Pa*LipA (yellow). **(B)** α-helical content per-residue averaged over three independent MD simulations for each of the three isolated domains MD1, EHD, and MD2. **(C)** The difference in the per-residue AHC of *Pa*Lif in complex with *Pa*LipA with respect to free *Pa*Lif computed over ten independent MD simulations. In (B) and (C), error bars indicate SEM. *p < 0.05 calculated by two-sided t-test.

In EHD, helices H6 - H8 and in MD2, H9 and H11 showed pronounced and repeated changes in the per-residue α-helix propensity, suggesting that in free *Pa*Lif, EHD and MD2 undergo helix-coil transitions. Second, I also computed the difference in the AHC per-residue of full-length *Pa*Lif (omitting residues 1-65) in complex with *Pa*LipA with respect to free *Pa*Lif (Figure 13C). A significant increase in AHC up to 54% was seen in MD2 (H9) when *Pa*Lif binds to *Pa*LipA, whereas in the EHD both increases and decreases of AHC are found. These

results suggest that MD2, and to some extent EHD, undergo conformational changes, and MD2 in particular gains secondary structure upon binding to *Pa*LipA.

Next, I studied the relative movements of these three domains. For this, I computed the dihedral angle defined by four residues in MD1 (Q137), EHD (A215, E268), and MD2 (R296) (Figure 14A), as a measure for a twisting motion of *Pa*Lif, and the radius of gyration (R_G), as a measure of *Pa*Lif's compactness.

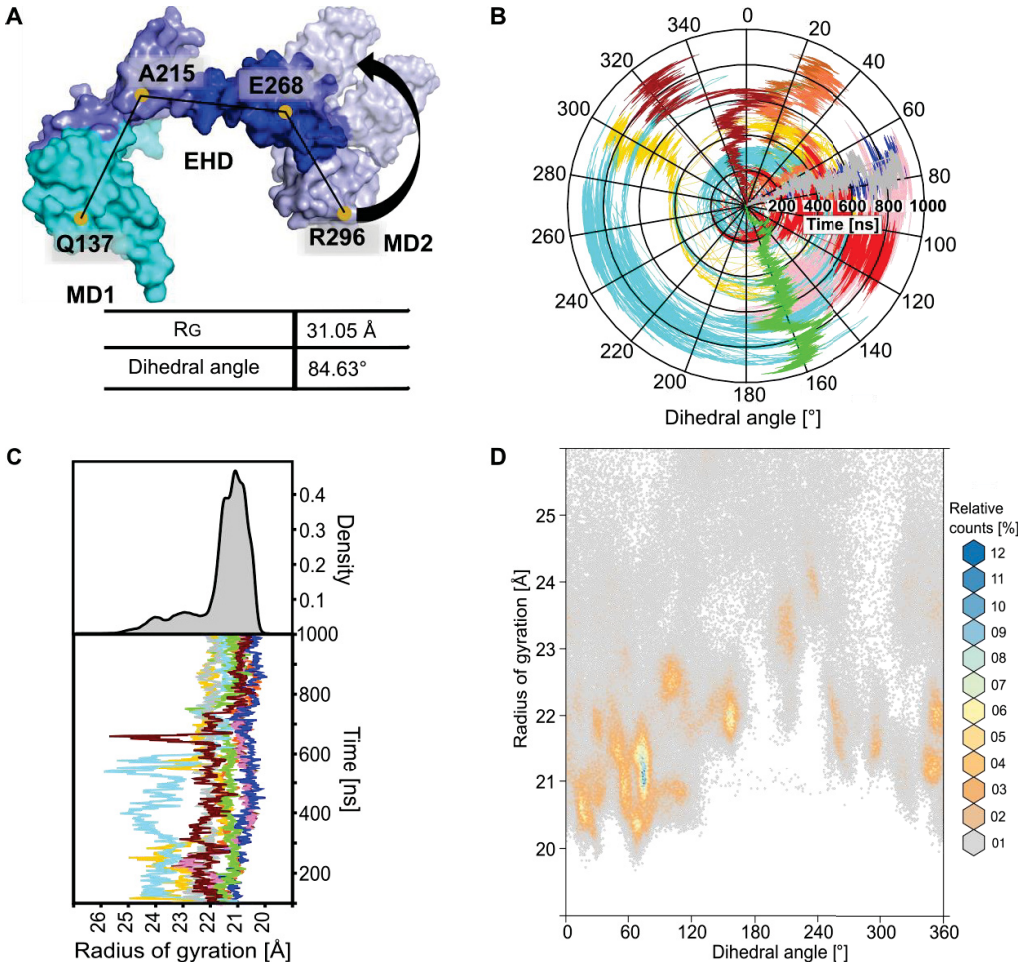


Figure 14. Structural dynamics of free *Pa*Lif. (A) Location of the residues (Q137, A215, E268, and R296) forming the dihedral angle that was used as a reaction coordinate to describe the twisting motions of free *Pa*Lif indicated by the black arrow. Domain coloring as in Figure 13A. The dihedral angle and R_G for bound *Pa*Lif are indicated below the structure. **(B)** Pronounced twisting motion (defined in panel A) of free *Pa*Lif is notable from a variation of dihedral angles as a function of simulation time (0 ns at the center, 1 µs at the rim). Each of ten independent MD simulations is indicated by a different color. **(C)** R_G of free *Pa*Lif as a function of time for ten independent MD simulations (different colors), of which the first 100 ns were discarded. On the top, the cumulative histogram is plotted. In all MD simulations, compaction of the structure was observed to an R_G of ~21 Å. **(D)** Distribution of conformations of free *Pa*Lif generated by MD simulations mapped to the R_G / dihedral angle plane. Clusters of highly abundant conformations (orange) among a vast number of low abundant conformations (gray) indicate the complexity of the structural dynamics of free *Pa*Lif.

Across the MD simulations, it was seen that *Pa*Lif adopts dihedral angles covering the entire range from 0° to 360° (Figure 14B), indicating pronounced twisting motions. However, the time traces showed, limited overlap of the sampled values among trajectories, which indicates that the twisting motions of free *Pa*Lif are slower than the timescale covered by the MD simulations. Additionally, R_G decreased to ~21 Å from 31 Å (R_G in bound conformation) indicating compaction of the structure in free form (Figure 14C). The compaction occurs on a time scale faster than a μs. Free *Pa*Lif likely can populate multiple conformations ranging from an extended to a compact state with different dihedral angles (Figure 14D).

4.2.2 Fluorescence analysis of *Pa*Lif

The work explained in this section was done by the co-authors of this publication Dr. Jakub Kubiak and Dr. Peter Dollinger.

Fluorescently labeled *Pa*Lif variants are active to *Pa*LipA

The effects of Cys mutations and covalently bound fluorescent dye on the function of *Pa*Lif were tested for *in vitro* efficiency. The six labeled double Cys *Pa*Lif variants were generated (named as P1-2, P1-3, P1-4, P2-3, P2-4 and P3-4) (Table 2) and tested for lipase activity assay. Labeled double Cys variants of *Pa*Lif revealed a comparable activation efficiency constant for all six *Pa*Lif variants.

Table 2. Fluorescence averaged dye distances $\langle R_{DA} \rangle_E$ [Å].

*Pa*Lif variant	Residue 1-2		$\langle R_{DA} \rangle_E$ [Å]		
			Homology model[a]	*Pa*Lif:LipA complex[b]	*Pa*Lif[b]
P1-2	Q137	A215	59	63 ± 5	58 ± 5
P1-3	Q137	E268	61	50 ± 7	39 ± 5
P1-4	Q137	R296	89	97 ± 7	60 ± 4
P2-3	A215	E268	51	66 ± 6	53 ± 6
P2-4	A215	R296	66	66 ± 6	48 ± 5
P3-4	E268	R296	53	52 ± 6	42 ± 7

[a] FRET averaged donor-acceptor distance simulated with FRET positioning and screening software (**section 2.6.4**) [191] using homology model of *Pa*Lif in complex with *Pa*LipA.

[b] FRET averaged donor-acceptor distance obtained with photon distribution analysis (PDA) [161, 192] of single-molecule multi-parameter fluorescence detection measurements.

Free *Pa*Lif undergoes conformational dynamics on the millisecond timescale

FRET experiments revealed that *Pa*Lif in solution is characterized by conformational dynamics on the sub-millisecond timescale indicated by the "dynamic shift" from the static FRET line observed in single-molecule multi-parameter fluorescence detection (smMFD) analysis [155].

Conformational heterogeneity of free *Pa*Lif in solution

To show that free *Pa*Lif in solution undergoes a dynamic exchange of states that occurs on a sub-millisecond time scale and is on average more compact than *Pa*Lif:LipA complex, an ensemble time-correlated single photon counting (eTCSPC) was done to obtain FRET-induced donor decay histograms and model donor-acceptor distance distributions $p(R_{DA})$ that can resolve conformational heterogeneity of highly dynamic molecules [157]. The analysis of all *Pa*Lif variants reveal broad $p(R_{DA})$ with mean inter-dye distance distributions $\langle R_{DA} \rangle$ (Table 2), whereas in the presence of micromolar *Pa*LipA, $p(R_{DA})$ shifts toward larger distance and is significantly narrower. This suggests an open and more restricted *Pa*Lif conformation with respect to structural flexibility, in complex with *Pa*LipA.

Next, fluorescence correlation techniques revealed that free *Pa*Lif exhibits more complex character, with anti-correlation of exchange between two states ranging from tens of nanoseconds, typical for unstructured peptide chain or dye linker diffusion [187, 188], through microseconds, characteristic for trapped states [189] and fast secondary structure elements folding [190], to hundreds of microseconds, signalizing high energy barriers [162]. All *Pa*Lif variants are similar except variant P3-4 (Table 2), which exhibits slower FRET dynamics with increased amplitude of sub-millisecond time scale. In order to exclude a possibility of experimental artifacts like fluorophore trapping at residue E268C, suggested by MD simulation and fluorescence experiments, four alternative *Pa*Lif variants ($P3^{255}$-4 (E255C-R296C), $P3^{258}$-4 (E258C-R296C), $P3^{259}$-4 (E259C-R296C) and $P3^{260}$-4 (E260C-R296C)) were tested. All four

alternative P3-4 variants show faster FRET dynamics, similar to dynamics of all other *Pa*Lif variants, except P3-4, but similar donor-acceptor distance distribution to P3-4 variant (Figure 15).

4.2.3 Structural ensemble of free *Pa*Lif

To gain insights at the atomistic level of the conformational variability of free *Pa*Lif, I combined the FRET data (**section 4.2.2**) with MD simulations. In order to evaluate the agreement between the computed conformational ensemble and that measured by FRET, we computed the distance distributions for the former (Figure 15). For the five distances P2-4, $P3^{255}$-4, $P3^{258}$-4, $P3^{259}$-4 and $P3^{260}$-4, the agreement with FRET data is good (χ^2 = 57.9, 49.9, 16.7, 15.3, and 6.4, respectively), and for the three distances P1-2, P1-4, P3-4, a fair agreement is found (χ^2 = 598.3, 128.6, and 463.0, respectively). The largest difference is seen for distances P1-3 and P2-3 (χ^2 = 2619.5 and 2670.7, respectively), which could be due to the high flexibility of the P3 region, as discussed in **section 4.2.2**. Still, the overall agreement of the conformational ensemble generated by unbiased MD simulations with experimental data is already remarkable.

To augment the data from the MD simulations with the information captured by FRET on long timescales, the simulated ensemble was reweighted using the MEM (**section 2.6.3**). The reweighted ensemble obtained at a relative entropy threshold corresponding to a free energy change of 3 *kT* (Figure 16A) is within the expected inaccuracy of MD force fields [193]. The distance distributions computed from the representative ensemble agree with the experimental data to within the estimated error ($\chi^2_{total} \approx 1.8$) (Figure 16A). As shown in Figure 16B, the optimal weights of many clusters obtained at the relative entropy threshold change by multiple orders of magnitude from their initial weights. Most (91 %) weights are down-weighted, while 9 % are up-weighted by at most a factor of 100. The representative ensemble contains about 35 distinct conformations of free *Pa*Lif that jointly contribute to the computed FRET distance distributions (**section 4.2.2**). These 35 conformations together constitute about 90 % of the

reweighted ensemble (Figure 16C). The largest cluster of the representative ensemble has a relative population of 15%. The 35 structures jointly cover the whole range of R_G values (**section 4.2.1**).

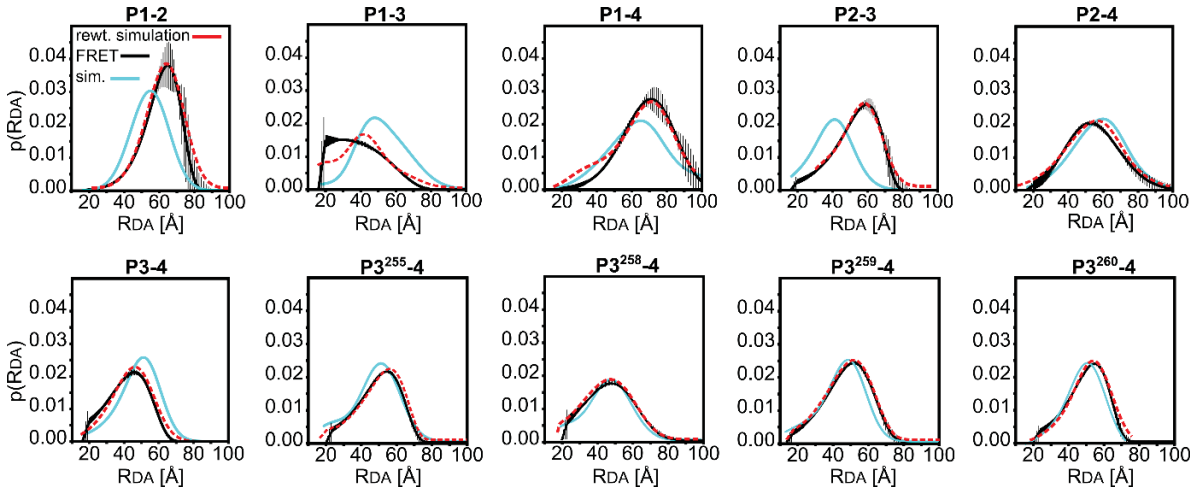


Figure 15. **Distance distribution profile obtained for free *Pa*Lif.** The probability of the average donor-acceptor distance ($p(R_{DA})$) is depicted as a function of the average donor-acceptor distance (R_{DA}) for all ten chosen residue pairs. Distance distributions obtained from FRET are shown in black with the error values as bars, distance distributions obtained from unbiased MD simulations (545 clusters) are shown in cyan color, and from the best refinement run in the red.

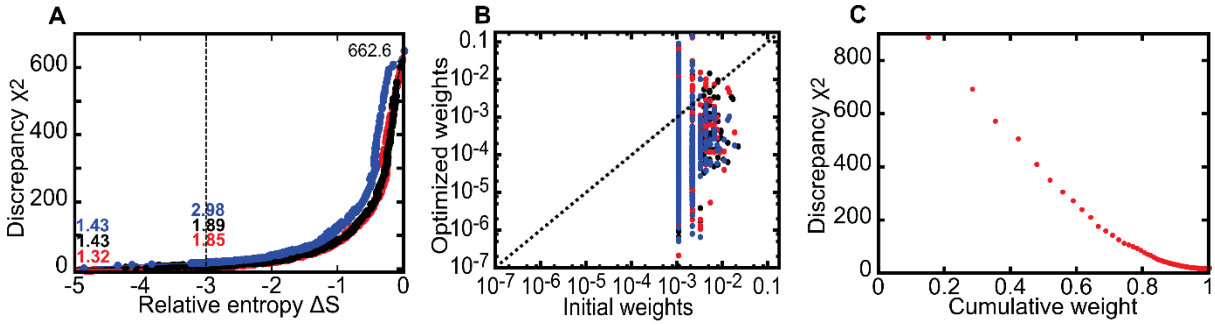


Figure 16. **Maximum entropy refinement of cluster weights for free *Pa*Lif. (A)** The discrepancy χ^2 between experimental and simulated data versus the relative entropy ΔS is depicted. The black-dotted line indicates the relative entropy threshold corresponding to a free energy change of 3 kT below which the simulated data is considered over-fitted. The blue, black, and red curves show three independent runs of refinement. **(B)** Change in initial cluster weights after optimization. The optimized weights are shown at the entropy threshold; the color code is as in panel A. **(C)** 35 most populated clusters after refinement. The discrepancy χ^2 as a function of the cumulative weight of the clusters is plotted for the best refinement run (red curve/dots in panels A/B). Clusters were rank-ordered by their optimized weights.

In order to examine the conformations of *Pa*Lif in the free state, a representative ensemble was further clustered using the R_G as a cluster criterion with a threshold value of 4 Å.

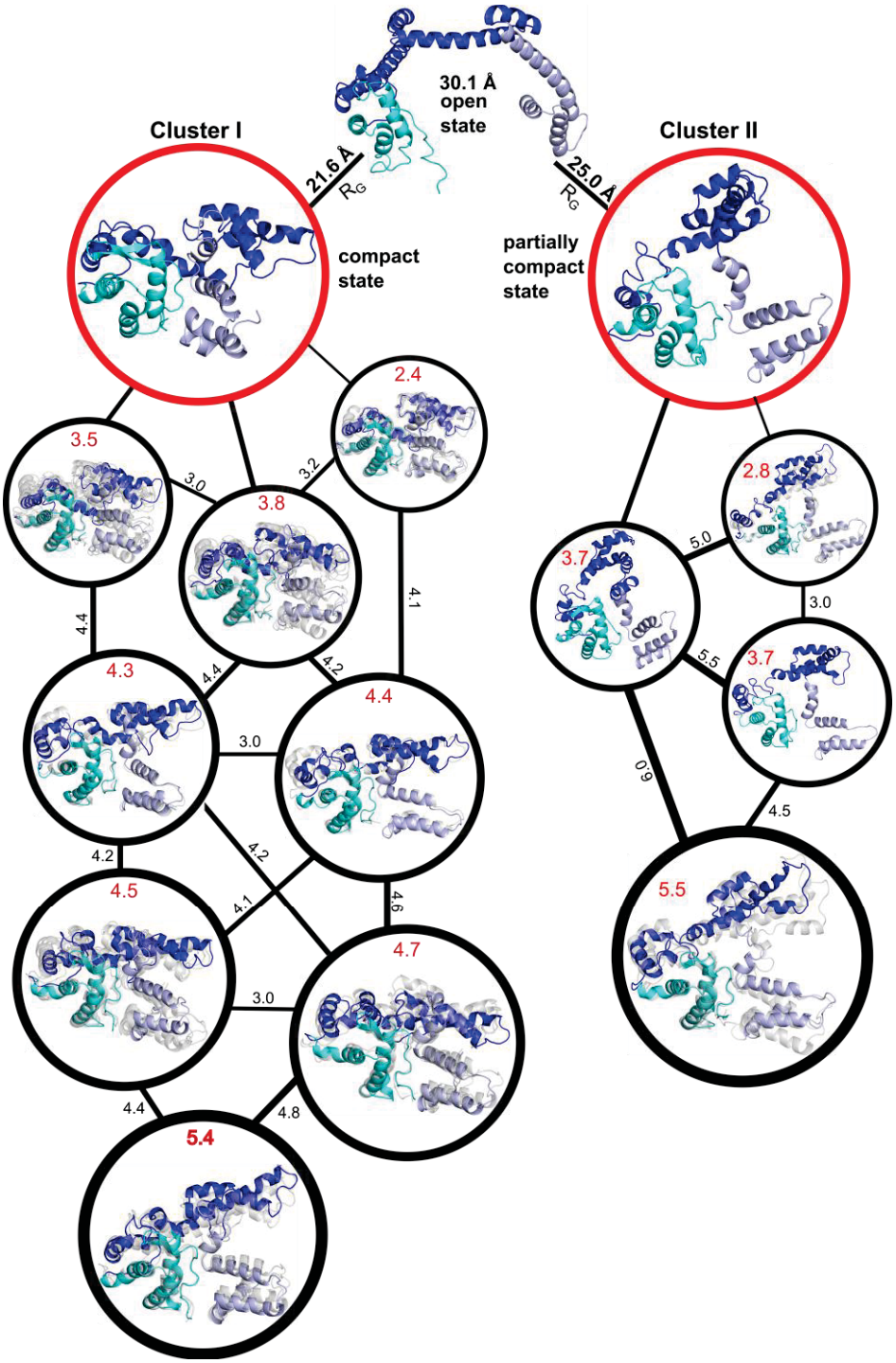


Figure 17. **Representative conformational ensemble of free *Pa*Lif.** The top 35 most populated clusters obtained after ensemble refinement by the MEM were first clustered with respect to the R_G with a cutoff of 4 Å; the two resulting clusters I and II (representative structures circled in red) were again clustered with respect to the C_α-RMSD with a cutoff of 3 Å. On the top, the structure of *Pa*Lif as found in the complex with *Pa*LipA is shown, with the R_G indicated below. The left branch depicts *Pa*Lif conformations of cluster I in the "compact states", with the thickness of the circles representing RMSD values of the sub-cluster relative to the representative structure; the RMSD values are also displayed in red within the circles. The right branch depicts *Pa*Lif conformations of cluster II in the "partially compact states". In addition to the representative conformation color-coded as in Figure 13A, further members of the sub-clusters are shown in gray. The thickness of the lines connecting the circles shows the C_α-RMSD between the sub-clusters; this value is also displayed beside the line.

The clustering showed that the representative ensemble is dominated by two clusters. Cluster I (27 structures) contains completely compact conformations of *Pa*Lif with an R_G of ~21 Å, whereas cluster II (8 structures) contains partially compact conformations of *Pa*Lif with an R_G of ~25 Å (Figure 17). We further analyzed the similarity of the structures within these two clusters, by sub-clustering them with respect to C_α-RMSD (threshold value of 3 Å). As shown in Figure 17, the structures were similar within the sub-clusters: For cluster I, eight sub-clusters were found, the farthest one having a C_α-RMSD of 5.4 Å from the centroid structure of the initial cluster; for cluster II, four sub-clusters were found, the farthest one having a C_α-RMSD of 5.5 Å from the centroid structure of the initial cluster.

4.2.4 Compact conformations of free *Pa*Lif hold affinity to bind *Pa*LipA

To further understand the nature of the compactness of free *Pa*Lif, and how such compact *Pa*Lif still retains the capability to interact with *Pa*LipA, we computed the probability density distribution of the total number of hydrogen bonds (HBs) and salt bridges (SBs) formed by *Pa*Lif in the free state as well as in complex with *Pa*LipA. For the latter, HBs and SBs were determined from conformations of ten independent MD simulations of 1 µs length each. For the former, the reweighted and representatives conformational ensembles of free *Pa*Lif were used, as was free *Pa*Lif in the bound conformation.

The total number of HBs and SBs formed by free *Pa*Lif in the bound conformation increased upon compaction of the free *Pa*Lif, which agrees well with the total number of HBs and SBs formed by *Pa*Lif in the complex (Figure 18A, B). Hence, these findings indicate that free, compact *Pa*Lif minimizes the number of unsatisfied polar interaction sites, resulting in similar numbers of interactions as in the complex. Changes in the SASA of all interface residues as well as hydrophobic residues are at variance with this (Figure 18C, D). While the distributions of free, compact *Pa*Lif show maxima at lower values than that of free *Pa*Lif in the bound conformation, the distributions only partially overlap with those computed for *Pa*Lif in

complex. This demonstrates that free, compact *Pa*Lif reduces the amount of overall interface SASA as well as hydrophobic SASA; yet, the remaining difference with respect to bound *Pa*Lif indicates that the *Pa*Lif:LipA complex should be able to gain affinity from a further burial of, particularly hydrophobic, SASA.

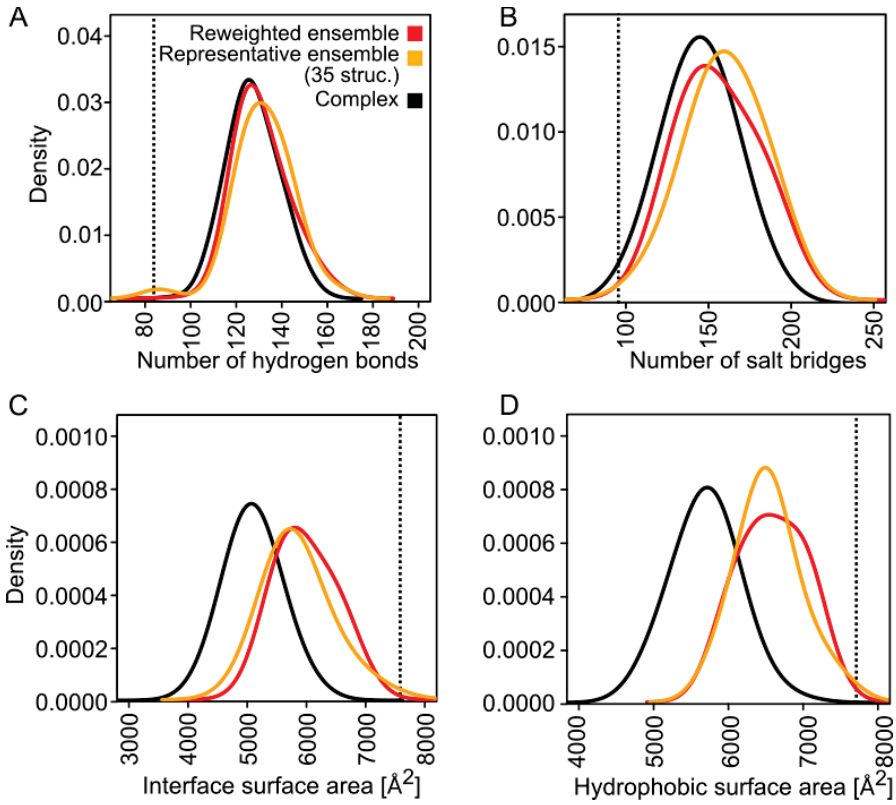


Figure 18. Probability density distributions of the number of total HBs, SBs, and SASA of all interface residues and hydrophobic residues for free *Pa*Lif and *Pa*Lif in complex with *Pa*LipA. (A) Probability density of the total number of HBs for the ensemble reweighted by MEM of free *Pa*Lif (red), the representative conformational ensemble of 35 structures (yellow), and *Pa*Lif in complex with *Pa*LipA (Black). The black-dotted line shows the value of the free *Pa*Lif in the bound conformation. The distributions of the number of HBs formed by conformations in the reweighted and representatives ensembles agree well with the distribution of HBs formed by *Pa*Lif in complex with *Pa*LipA. **(B)** Probability density of the total number of SBs; the color code is as in panel A. Again, all three distributions agree well with each other. **(C)** Probability density distribution of the SASA of interface residues; the color code is as in panel A. The distributions of SASA for all interface residues of the respective ensembles partially overlap with the distribution of bound *Pa*Lif. **(D)** Probability density distribution of the SASA of hydrophobic residues; the color code is as in panel A. The distributions of SASA of hydrophobic residues of the respective ensembles partially overlap with the distribution of bound *Pa*Lif.

4.3 Conclusion and significance

In this study, the structural stability of the individual domains of free *Pa*Lif were analyzed in free as well as in the bound conformation with *Pa*LipA. The results revealed that separated MD1 retained almost completely the α-helical secondary structure, despite the presence or

absence of *Pa*LipA. The combination of MD simulations with the experimental characterization by FRET observables revealed the complex dynamics of structurally minimalistic free *Pa*Lif on the timescale of milliseconds.

The main results of this study are:

- We showed that the two domains, MD2 and EHD, undergo a helix-loop transition in free *Pa*Lif and gain higher helicality upon binding to *Pa*LipA. This finding indicates that both domains go through conformational rearrangement upon *Pa*LipA binding.
- The considerably stable secondary structure of MD1 in both free and bound conformations of *Pa*Lif led us to conclude that MD1 may provide a platform for the stable binding of *Pa*LipA in the complex.
- Dihedral angle, R_G, and FRET distances indicate that free *Pa*Lif attains multiple conformations. However, an inclination towards the compaction of the structure was observed in the free form compared to the "head-phone" like state in the complex with *Pa*LipA.
- Free *Pa*Lif undergoes conformational dynamics on a timescale of sub-milliseconds with reversibly exploring compact and open conformations.
- Our simulation results after reweighting produced atomistic structures of free *Pa*Lif that are consistent with measured FRET distance distributions.
- Despite the compact conformations, free *Pa*Lif retains the affinity to bind *Pa*LipA via partially exposed hydrophobic and interface residues. As not all hydrophobic and interface residues get buried upon compaction, they provide the gain of affinity to form complex with *Pa*LipA.

Thus, this study provides important insights into the structural-functional dynamics of free *Pa*Lif. Therefore, in the next chapter (**publication II**) we solved the solution NMR structure of structurally stable MD1 and investigated its role in binding and *Pa*LipA activation.

5. PUBLICATION II - STRUCTURAL AND DYNAMICAL INSIGHTS REVEALING HOW LIPASE BINDING DOMAIN MD1 OF *PSEUDOMONAS AERUGINOSA* FOLDASE AFFECTS LIPASE ACTIVATION

Aldino Viegas[$,#], Peter Dollinger[#], Neha Verma[#], Jakub Kubiak, Thibault Viennet, Claus A.M. Seidel, Holger Gohlke, Manuel Etzkorn, Filip Kovacic, and Karl-Erich Jaeger

[#] These authors contributed equally to this work.

Original publication, see pages 162-209

5.1 Background

We have shown in **publication I** that the MD1 of *Pa*Lif has a stable secondary structure in free as well as in bound conformations, which led us to speculate about its role in the binding of *Pa*LipA. The sequence alignment of *Pa*Lif and *Bg*Lif revealed that among the five domains, MD1 shares the highest sequence similarity (52%). The importance of MD1 for foldase activity was further highlighted by the finding that MD1 comprises the foldase sequence motif (**section 2.4**), which is evolutionarily conserved among all foldase families and upon mutation inactivates the foldase [113].

In this publication, the role of MD1 in the activation of *Pa*LipA has been investigated by biochemical analysis, NMR spectroscopy, fluorescence spectroscopy, and MD simulations. Particularly, the Y99A mutant was investigated. Y99 is an evolutionarily conserved amino acid in helix α1 of MD1. Upon mutation to alanine, the ability of *Pa*Lif to activate *Pa*LipA is lost (**section 2.4**), which is further investigated with regards to the mechanism of *Pa*LipA activation. For the first time, the NMR solution structures of MD1 and variant $MD1_{Y99A}$ have been solved and deposited in the protein data bank (PDB). I performed MD simulations of the comparative

model of the *Pa*Lif:LipA complex in connection with rigidity analyses which suggest a long-range network of interactions in the tight complex spanning from Y99 of *Pa*Lif to the active site of *Pa*LipA, which might be essential for its activation. These findings provide important details about the putative mechanism for *Pa*LipA activation.

5.2 Results

For the following sections (**section 5.2.1-5.2.4**), the text and figures have been taken and modified from **publication II**. The work explained in **sections 5.2.1- 5.2.3** was done by the co-authors of this publication and the work done by myself is presented in **sections 5.2.3 - 5.2.4**.

5.2.1 Isolated MD1, but not MD1$_{Y99A}$ decelerates *Pa*Lif-induced activation of *Pa*LipA

It was proposed that specific interactions of MD1 with *Bg*LipA are important for its activation in *B. glumae* [19, 80, 111]. Our *in vitro* experiments showed that the isolated MD1 could not activate pre-active *Pa*LipA either. However, the addition of MD1 to pre-active *Pa*LipA in 12 to 20-fold molar excess during activation of *Pa*LipA with *Pa*Lif significantly slowed down the activation reaction. This result indicates that isolated MD1 can interfere with *Pa*Lif's capability to activate *Pa*LipA. This effect was not observed with isolated MD1$_{Y99A}$. Additionally, a stronger binding affinity was observed for *Pa*Lif (K_D = 39 ± 13 nM) compared to *Pa*Lif$_{Y99A}$ (K_D = 102 ± 35 nM).

5.2.2 *Pa*LipA complexes with *Pa*Lif and variant *Pa*Lif$_{Y99A}$ exhibit similar unfolding profiles

During thermal unfolding analyzed by the intrinsic protein fluorescence of respective solutions, *Pa*Lif and *Pa*Lif$_{Y99A}$ alone show typical unfolding curves with an unfolding temperature of ~50°C. The thermal unfolding curve of pre-active *Pa*LipA does not show a sharp unfolding transition typical for folded proteins but rather a broad transition with a maximum at 73°C. Addition of MD1 or MD1$_{Y99A}$ to pre-active *Pa*LipA did not considerably affect this unfolding

curve, indicating that the domains do not change the fold of pre-active *Pa*LipA and do not strongly interact with *Pa*LipA, which was also confirmed by fluorescence binding assay. In contrast, the addition of *Pa*Lif or *Pa*Lif$_{Y99A}$ to pre-active *Pa*LipA yielded typical unfolding curves with a maximum of the first derivatives at ~37°C.

Fluorescence measurements show that *Pa*Lif and *Pa*Lif$_{Y99A}$ exhibit similar hydrodynamic properties (rotational correlation time global ρ_{global} = 33 ns) as well high degrees of internal flexibility of the protein in the absence of *Pa*LipA (order parameter S^2 = 0.3). In the case of *Pa*Lif:LipA complex, internal flexibility of *Pa*Lif is decreased (S^2 = 0.38), indicating that *Pa*Lif is more rigid in bound form. However, for *Pa*Lif$_{Y99A}$:LipA, internal flexibility remains high (S^2 = 0.22). This finding indicates that for the mutant the MD1 domain does not (or weakly) interact with *Pa*LipA while the rest of *Pa*Lif still interacts normally.

5.2.3 Structural insights into MD1 of *Pa*Lif and variant *Pa*Lif$_{Y99A}$

To obtain the first structural insights into this system and to investigate the effects of the critical Y99A mutation on the structure of the MD1 domain, we here solved the NMR solution structure of the isolated MD1 domain (Figure 19A, PDB code: 5OVM; BMRB code 34175) as well as of MD1$_{Y99A}$ variant (Figure 19B, PDB code: 6GSF; BMRB code 34286).

Both MD1 and MD1$_{Y99A}$ resemble a three α-helical bundle preceded by 27 N-terminal residues without clear secondary structure with an RMSD$_{Cα}$ of 2.4 Å, when comparing the MD1 and MD1$_{Y99A}$ structural ensembles. Only minor structural differences were observed within each ensemble of 20 energetically most favorable structures for MD1 as well as for MD1$_{Y99A}$, as indicated by RMSD$_{Cα}$ of 1.3 ± 0.3 Å and 0.8 ± 0.2 Å, respectively. The obtained structures of the isolated MD1 variants are similar to the respective domain in the crystal structure of the *Bg*Lif:LipA complex (Figure 19C) [19] showing that this domain adopts a stable fold, even when isolated, and in the absence of a lipase.

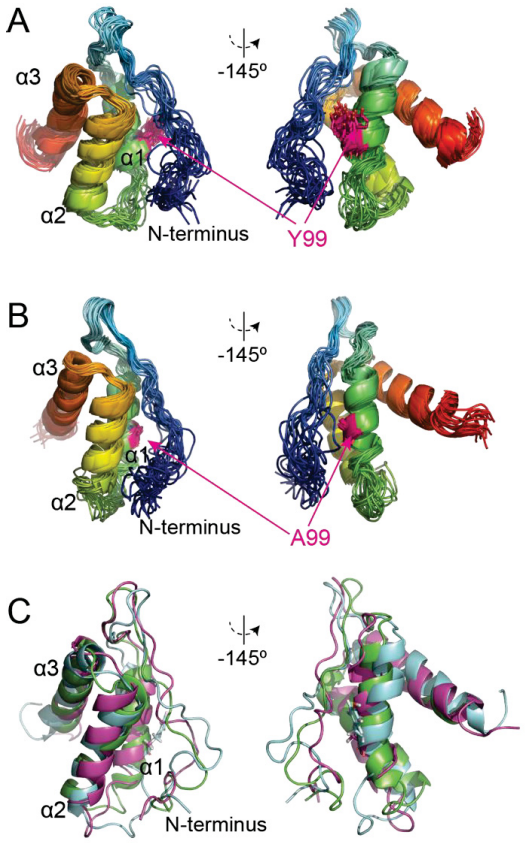


Figure 19. Details of MD1 and variant MD1$_{Y99A}$ structures obtained by NMR spectroscopy. (A) Cartoon representations of the structural ensemble of the 20 best solution structures of MD1 and **(B)** MD1$_{Y99A}$ variant. **(C)** Comparison of the representative NMR solution structures of MD1 (cyan) and MD1$_{Y99A}$ (purple) with the crystal structure of MD1 from *B. glumae* (green) (PDB code: 2ES4) [19].

5.2.4 *Pa*Lif$_{Y99A}$ exerts a long-range effect on *Pa*LipA which may destabilize the structure of the substrate binding pocket

To understand the possible role of the mutation Y99A for *Pa*Lif-induced activation of *Pa*LipA, we initially compared the X-ray structure of *Pa*LipA$_o$ in the active conformation (PDB code: 1EX9) with the X-ray structures of *Bg*LipA$_c$ in the inactive conformation, one in the complex with its specific foldase (PDB code: 2ES4) and the other in the unbound form (PDB code: 1QGE) (Figures. 20A-C). As discussed in **section 2.4,** in the open conformation of lipase A, helix α5 is moved away from the active site, and a short two-stranded β-sheet close to the active site is formed by residues 21-22 and 25-26. Residues 17-30 shape the substrate binding pocket [17], which indicates the relevance of this short two-stranded β-sheet structure in the region of

residues 17-30 for *Pa*LipA activation. Because this β-sheet is not formed in the unbound conformation of *Bg*LipA$_c$, where helix α5 is occluding the active site, we postulated this structural element as a hallmark of the open and active lipase A. Notably, *Bg*LipA in the complex, has an overall structure identical to unbound *Bg*LipA$_c$, but does have a two-stranded β-sheet in the region of residues 17-30, yet helix α5 is in the closed conformation. The structural comparison thus indicates that the foldase induces the formation of the two-stranded β-sheet during activation of the lipase A. Hence, the foldase-bound lipase A can be considered an intermediate conformation on its way to an open conformation, with residues 17-30 acting as a "loaded spring" (**section 2.4**).

The influence of the Y99A mutation on the structural stability of *Pa*LipA was investigated by using an ensemble-based perturbation approach [194] integrated into the CNA approach (**section 2.6.6**). CNA was applied on the conformational ensemble of the *Pa*Lif:LipA complex generated from the above MD simulations. We followed that approach because *Pa*LipA strongly binds to both *Pa*Lif variants, suggesting that the respective complex structures are very similar in both cases. The computed changes in the residue-wise free energy $\Delta G_{i,\mathrm{CNA}}$ (**section 2.6.6**), a measure for structural stability [194], was largest for residues 1-45, 197-202, 242-250, and 268-286 in *Pa*LipA (Figures 20E and 20F).

Notably, these affected residues form a narrow pathway that reaches the β-sheet-forming region including residues 28-30 (Figures 20E, 20F), indicating that the mutation decreases the stability of substrate binding pocket (Figure 20D). We speculate that this decrease in stability prevents the partial β-sheet formation of substrate binding pocket and, thus, disfavors *Pa*LipA activation by *Pa*Lif$_{Y99A}$. Finally, the Y99A mutation also affects the stability of a number of residues in *Pa*Lif itself, especially the neighboring residues 66-80 and 89-115 in MD1 (Figure 20G), this agrees with the minor changes seen in the NMR structures of MD1 and MD1$_{Y99A}$ variant (**section 5.2.3**).

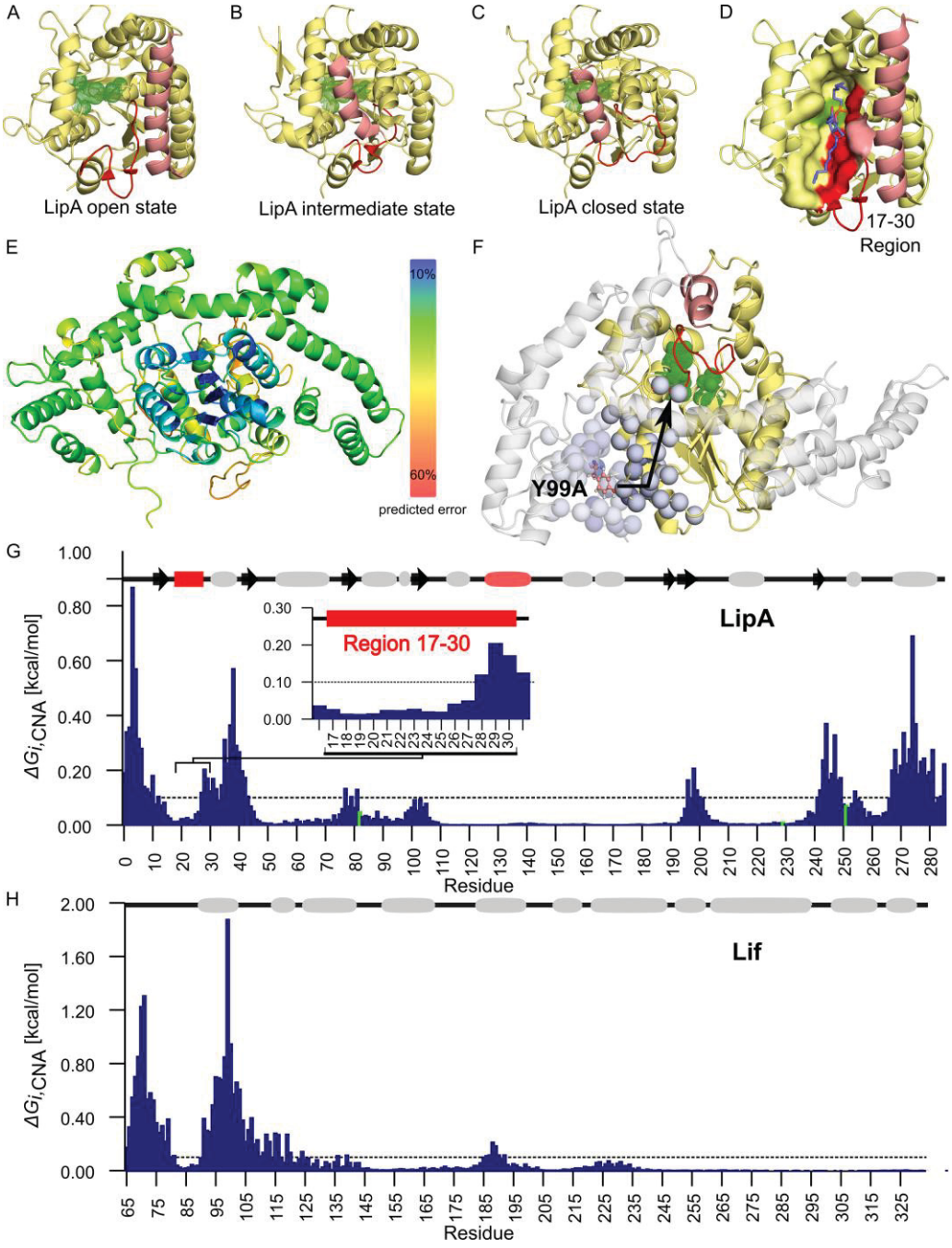


Figure 20. Influence of mutation Y99A in *Pa*Lif on the structural stability of *Pa*LipA. (A) Structure of *Pa*LipA with OCP inhibitor bound in the active site crystallized in the open conformation (PDB code: 1EX9) [17], in which the helix α5 (salmon) is moved away from the active site (green). A short two-stranded β-sheet close to the active site is formed by residues 17-30 (red). **(B)** *Bg*LipA from the crystal structure of the *Bg*Lif:LipA complex (PDB code: 2ES4) [19]. The lipase shows a two-stranded β-sheet (red), a characteristic feature of the open (active) conformation, nevertheless helix α5 (salmon) adopts a closed (inactive) conformation. This suggests a foldase-induced formation of a two-stranded β-sheet during activation of the lipase. **(C)** Crystal structure of *B. glumae* lipase crystallized in the closed conformation (PDB code: 1QGE) [195] with helix α5 (salmon) covering the active site (residues as in panel A). In this conformation, a two-stranded β-sheet close to the active site is not formed. Residues 17-30 of *B. glumae* lipase, forming a two-stranded β-sheet, are indicated in red. **(D)** Crystal structure of active *Pa*LipA with inhibitor OCP bound in the active site. The region of residues 17-30 forms part of the active site (red), required for the binding of the ligand. **(E)** CNA was applied on an ensemble of structures of the *Pa*Lif:LipA complex generated from ten independent MD simulations. Residues with $\Delta G_{i,\mathrm{CNA}}$ above a threshold

of 0.1 kcal mol^{-1} are depicted as spheres on the *Pa*Lif:LipA complex structure. Blue colors reflect predicted $\Delta G_{i,\mathrm{CNA}}$ values; the larger the value, the darker is the color. The black arrow indicates how the perturbation by Y99A mutation of *Pa*Lif (pink, ball and stick representation) influences residues in *Pa*LipA. Due to the decrease in the stability of the surrounding region of residues 17-30 in *Pa*LipA, we speculate that the conformational changes required for the intermediate state of *Pa*LipA on the way of activation is hampered upon *Pa*Lif$_{Y99A}$ mutation. The color code for helix α5, residues 17-30, and the active site is as in panel A. **(F)** The histogram shows the per-residue $\Delta G_{i,\mathrm{CNA}}$ for *Pa*LipA. The dashed line at 0.1 kcal mol^{-1} indicates the threshold above which residues are considered perturbed. The standard error of the mean is < 0.05 kcal mol^{-1} for all residues. **(G)** Per-residue $\Delta G_{i,\mathrm{CNA}}$ shown for *Pa*Lif, with the same threshold. The standard error of the mean is < 0.05 kcal mol^{-1} for all residues.

5.3 Conclusion and significance

In this publication, we showed that the Y99A mutation in MD1 induces only minor changes in the domain's overall structure. However, activation of *Pa*LipA induced by *Pa*Lif is inhibited by addition of MD1, but not by the addition of its variant MD1$_{Y99A}$. CNA suggested a network of long-range interactions that span from Y99 of *Pa*Lif to the active site of *Pa*LipA. This network supports the formation of a key secondary structure element (a partial β-sheet) in residues 17-30 of *Pa*LipA on the pathway from the pre-active to the active state.

Further results of the study are:

- The *Pa*Lif$_{Y99A}$ variant, which carries a single amino acid mutation in the conserved foldase motif, did not activate *Pa*LipA, in contrast to *Pa*Lif.
- Previously, Shibata and co-workers reported that *Pa*Lif$_{Y99C}$, *Pa*Lif$_{Y99H}$, *Pa*Lif$_{S102R}$, and *Pa*Lif$_{R115C}$ variants (all carrying mutations in the foldase conserved motif) do not form complexes with *Pa*LipA [111] (**section 2.5**). By contrast, we showed that *Pa*Lif$_{Y99A}$ binds pre-active *Pa*LipA.
- *Pa*Lif binding to pre-active *Pa*LipA was found to be slightly stronger than to *Pa*Lif$_{Y99A}$.
- The MD1 solution structure provides the first experimental evidence that this domain adopts a stable tertiary structure even in the absence of *Pa*LipA.

Thus, our study for the first time provides insights at the atomistic level as to a potential mechanism of *Pa*Lif-mediated pre-active-to-active *Pa*LipA folding. This finding is further investigated in **publication III** in order to validate this putative mechanism.

6. PUBLICATION III - THE MEMBRANE-INTEGRATED STERIC CHAPERONE LIF FACILITATES ACTIVE SITE OPENING OF *PSEUDOMONAS AERUGINOSA* LIPASE

Neha Verma[1], Peter Dollinger[2], Filip Kovacic[2], Karl-Erich Jaeger[2,3], and Holger Gohlke[1,4]

Original publication, see pages 210-242

6.1 Background

In the previous **publication II**, we showed that the network of long-range interactions formed by Y99 of *Pa*Lif supports the formation of the partial β-sheet structure of residues 17-30 in the binding site of *Pa*LipA. As discussed in **section 5.2.4** of **publication II**, this partial β-sheet structure is a characteristic feature of the open, active lipase A, but absent in the closed, inactive lipase A (Figure 20A). As this partial β-sheet structure was also present in the crystal structure of the *Bg*Lif:LipA complex, where the lid is still in the closed conformation, we hypothesize that the foldase-bound *Bg*LipA is in an intermediate conformation where the partial β-sheet formation has occurred but not yet the lid opening, comparable to a "loaded spring" ready to change to the open conformation. In the crystal structure of *Pa*LipA$_o$ (PDB code: 1EX9), the partial β-sheet structure in the region of residues 17-30 forms a part of the binding site and likely stabilizing the neighboring loops (Figure 20B-D).

To this point, the complete mechanism of *Pa*LipA activation mediated by *Pa*Lif at the atomistic level is still not known. Hence, in this publication, I evaluated the above stated hypothesis and proposed the activation mechanism of *Pa*LipA via *Pa*Lif by using unbiased and biased MD simulations followed by configurational free energy computations, complemented by *in vitro* biochemical experiments for validation. The biochemical experiments were performed by Dr. Peter Dollinger.

6.2 Results

For the following sections (**section 6.2.1-6.2.4**), the text and figures have been taken from **publication III**. In this study, I performed biased and unbiased molecular simulations, CNA, analyzed the data and explained my work **in sections 6.2.1, 6.2.2** and **6.2.4**.

6.2.1 Structural dynamics of the lid of free *Pa*LipA and when bound to *Pa*Lif

Initially, we analyzed the tendency of *Pa*LipA$_c$ free and in the complex *Pa*Lif:LipA to move towards the open state, and of free open *Pa*LipA$_o$ towards the closed state, by unbiased MD simulations. Due to the absence of respective crystal structures, *Pa*LipA$_c$ and *Pa*Lif:LipA were built by homology modeling (Figures 21A and 21B).

First, we analyzed the MD simulations with respect to the average β-sheet propensity of residues 17-30 of *Pa*LipA because this secondary structure type is a characteristic feature of *Pa*LipA$_o$ (**section 5.2.4 and 6.1**). For the *Pa*Lif:LipA complex, the likelihood of β-sheet formation is highest (96.5 ± 0.8%, mean ± SEM) (Figure 21D). In contrast, *Pa*LipA$_c$ showed a significantly lower β-sheet propensity of 41.1 ± 12%. This result indicates that *Pa*Lif fosters the formation of the β-sheet structure. As expected, *Pa*LipA$_o$ exhibits a β-sheet propensity more similar to that of the *Pa*Lif:LipA complex (83.5 ± 3.5%), yet, the significantly smaller value suggests that *Pa*LipA$_o$ tends to move towards the closed conformation.

Next, we computed the distance between the centers of mass (D_{COM}) of helix α5 and α8 over the MD simulations to measure the opening and closing of the active site. Starting from *Pa*LipA$_c$, pronounced fluctuations of D_{COM} were observed that encompass both partially open lid conformations ($D_{COM} \approx 16$ Å) and more closed ones ($D_{COM} \approx 10$ Å) compared to the starting state ($D_{COM} = 13.3$ Å) (Figure 21E). A similar behavior was observed for *B. cepacia* lipase A during MD simulations in water [196].

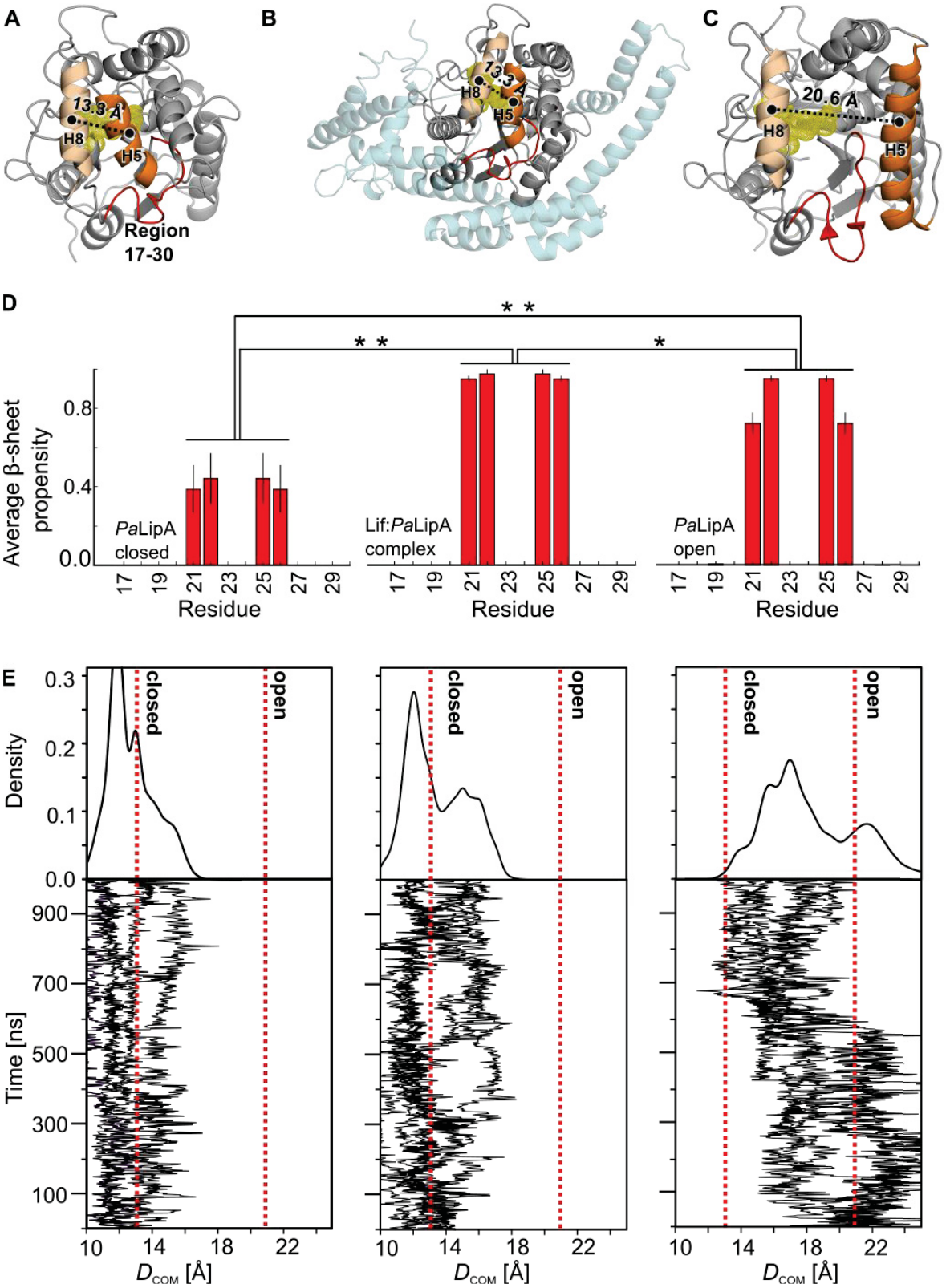


Figure 21. Structural dynamics of *Pa*LipA during unbiased MD simulations. (A) Structural model of *Pa*LipA$_c$ generated by homology modeling. Active site residues (catalytic triad residues S82, H251, and D229) are shown as yellow dots, which are covered by the lid domain (helix α5) (orange). Helix α8 is highlighted in wheat. D_{COM} is represented by a black, dotted line. In *Pa*LipA$_c$, D_{COM} = 13.3 Å. Residues 17-30 are shown in red and do not exhibit the partial β-sheet structure, which is a characteristic feature of *Pa*LipA$_o$. **(B)** Homology model of the closed *Pa*LipA in complex with *Pa*Lif, *Pa*Lif is represented in cyan, otherwise the representation is as in (A). **(C)** Crystal

structure of *Pa*LipA$_o$ (PDB code: 1EX9), represented as in (A). Helix α5 moved away from the active site, and D_{COM} = 20.6 Å (black-dotted line). Residues 17-30 form a partial β-sheet structure (red). **(D)** Average per-residue β-sheet propensities of residues 17-30 starting from *Pa*LipA$_c$ (left), *Pa*Lif:LipA (middle), and *Pa*LipA$_o$ (right). Error bars indicate SEM and stars statistically significant differences of the averages calculated with the Student's *t*-test:"*" *p*-values < 0.05 and "**" *p*-values < 0.001. **(E)** D_{COM} over the simulation times of six MD trajectories each for the three systems listed in (D). Additionally, the probability densities are shown. Red-dotted lines indicate the D_{COM} values of the open and closed *Pa*LipA states.

Starting with the *Pa*Lif:LipA complex, the probability density of partially open conformations (D_{COM} ≈ 16 Å) was ~2-fold higher than for *Pa*LipA$_c$ (Figure 21E). This suggests that *Pa*LipA in complex with *Pa*Lif has a stronger tendency to move toward a partially open state than *Pa*LipA$_c$, although this tendency is obvious in only three trajectories out of six. As when starting from the closed conformation of *Pa*LipA, further closing of the lid was also observed during the MD simulations of the complex. Finally, starting from *Pa*LipA$_o$, the partially open state becomes most populated (D_{COM} ≈ 16 Å), and even closed conformations (D_{COM} ≤ 13.3 Å) were found (Figure 21E). Likewise, for *B. cepacia* lipase A, an open-to-closed transition of the lid during MD simulations in water was found [197].

6.2.2 The open state of *Pa*LipA is more favorable when bound to *Pa*Lif

To complement the unbiased MD simulations, we computed the PMF of the opening of the active site in free *Pa*LipA and in complex with *Pa*Lif, applying US and using D_{COM} (Figure 21A) as a reaction coordinate. The PMF computations were performed for a plausible transition path of helix α5 obtained from unbiased MD simulations of *Pa*LipA$_o$. For comparison, the PMF values at the smallest D_{COM} sampled (11.6 Å) were set to zero in both cases (Figure 22A).

Although in both cases the configurational free energy increases with increasing D_{COM}, the PMFs differ in their global shape: The PMF of the *Pa*Lif:LipA complex increases more moderately than that of *Pa*LipA and shows broader local minima (Figure 22A). In more detail, the global energy minima (state I) for both free *Pa*LipA and the *Pa*Lif:LipA complex are found for the closed state (D_{COM} = 12.4 Å and 13.2 Å, respectively, $\Delta G \approx 0$ kcal mol^{-1} with respect to

D_{COM} = 11.6 Å). At D_{COM} ≈ 14.8 Å, both PMFs have a local minimum (state II) of similar height (ΔG ≈ 1 kcal mol^{-1}). The corresponding energy well of the *Pa*Lif:LipA complex is extended until D_{COM} ≈ 16 Å. In contrast, the PMF for *Pa*LipA rises steeply immediately following the local minimum. Finally, flat PMF regions are found for both systems at D_{COM} ≈ 20.6 Å (state III), but the configurational free energies with respect to the global minimum differ (*Pa*Lif:LipA: ΔG ≈ 2.9 kcal mol^{-1}, *Pa*LipA: ΔG ≈ 4.6 kcal mol^{-1}).

Furthermore, we computed the average β-sheet propensity of residues 17-30 of *Pa*LipA with and without *Pa*Lif over the reweighted configurations from umbrella sampling for states I-III, respectively (Figure 22B). At the global minimum (state I), the β-sheet propensity averaged over windows 1 and 2 is significantly lower for *Pa*LipA (~21 ± 6%) than in state II, averaged over windows 3 and 4, and state III, averaged over windows 9 and 10 (~ 60 ± 4% and ~ 63 ± 3%, respectively). In contrast to *Pa*LipA, in *Pa*Lif:LipA, state I (~ 94 ± 0.8%) has a similar β-sheet propensity as state II and state III (~ 75 ± 2% and 84 ± 1%, respectively). The difference between the average β-sheet propensities of states I of *Pa*LipA and *Pa*Lif:LipA is highly statistically significant ($p < 0.001$). Similarly, for states II and III, the respective average β-sheet propensities of *Pa*Lif:LipA are significantly higher than those of *Pa*LipA ($p < 0.05$ for both states).

To conclude, the PMF computations reveal that the open state of *Pa*LipA is disfavored compared to the closed state but that in *Pa*Lif:LipA the open state is 1.7 kcal mol^{-1} more favorable than in *Pa*LipA. Furthermore, according to unbiased configurations from the umbrella sampling simulations, binding to *Pa*Lif significantly favors the formation of the β-sheet in the region of residues 17-30, and this effect is most pronounced in the state I (~73 fold increase in the propensity).

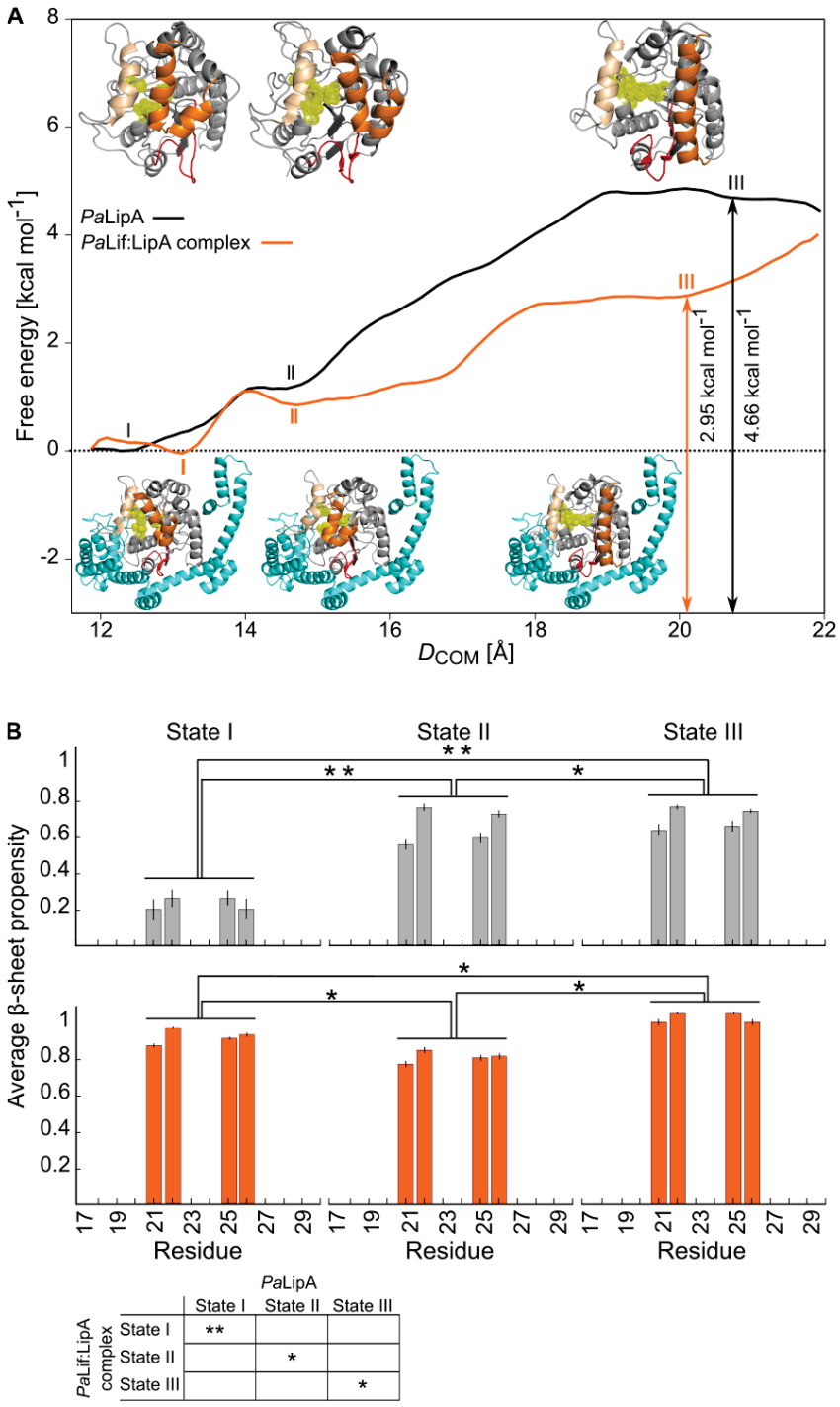


Figure 22. PMF computation of the active site opening in *Pa*LipA and the *Pa*Lif:LipA complex and average β-sheet propensities of residues 17-30 of *Pa*LipA with and without *Pa*Lif for three states identified in the PMFs. (A) Configurational free energies of active site opening of *Pa*LipA as a function of D_{COM} used as a reaction coordinate for free *Pa*LipA (black) and the *Pa*Lif:LipA complex (orange). The standard deviation for all data points is < 0.002 kcal mol^{-1} computed by bootstrap analysis. Roman numbers indicate the identified states. Representative structures for states I-III are shown as cartoons for *Pa*LipA (top) and *Pa*Lif:LipA complex (bottom), respectively. The PMF values at D_{COM} = 11.6 Å were set to zero, respectively. **(B)** Per-residue averaged β-sheet propensity for residues 17-30 of *Pa*LipA, calculated across the US windows corresponding to the states I-III as described in (A), using reweighted (unbiased) configurations for *Pa*LipA (top) and the *Pa*Lif:LipA complex (bottom). The table at the bottom displays results from comparing β-sheet propensities between *Pa*LipA and *Pa*Lif:LipA. Error bars indicate the SEM and statistically significant differences calculated with the Student's *t*-test were indicated by "*" *p*-values < 0.05 and "**" *p*-values < 0.001.

6.2.3 *Pa*LipA released from *Pa*Lif loses its lipolytic activity over time under *in vitro* conditions

To validate the above computational finding, biochemical experiments were performed in which a catalytically inactive *Pa*LipA variant, in which amino acid S82 of the catalytic triad is mutated to alanine, was used in addition to wild type *Pa*LipA. Purified *Pa*LipA$_{S82A}$ was renatured and used for complex formation with *Pa*Lif at 1 μM concentration. According to the results of thermal unfolding experiments carried out with differential scanning fluorimetry, *Pa*LipA$_{S82A}$ forms a complex with *Pa*Lif that has stability similar to that of *Pa*LipA with *Pa*Lif. As expected, no catalytic activity is found for the *Pa*Lif:LipA$_{S82A}$ complex, in contrast to the *Pa*Lif:LipA complex. After dilution of the *Pa*Lif:LipA$_{S82A}$ complex to 1 nM, renatured *Pa*LipA was added in excess at a concentration of 100 nM, followed by three h incubation. The addition of *Pa*LipA to the *Pa*Lif:LipA$_{S82A}$ complex restored activity to ~90 % of that of *Pa*Lif:LipA, indicating that *Pa*LipA replaces *Pa*LipA$_{S82A}$ and then becomes activated by *Pa*Lif.

Finally, a complementary experiment was performed in which 50 nM *Pa*Lif:LipA complex was supplemented with 50 nM or 100 nM *Pa*LipA$_{S82A}$, and with buffer as control, followed by determination of the catalytic activity over time. After 145 min, the activity level decreased by about 15 % and 35 % in the presence of 50 nM and 100 nM *Pa*LipA$_{S82A}$, respectively, which indicates the replacement of *Pa*LipA by *Pa*LipA$_{S82A}$ and the subsequent loss of catalytic activity of free *Pa*LipA in a *Pa*LipA$_{S82A}$ concentration-dependent manner. Addition of 100 nM *Pa*Lif after 150 minutes restored catalytic activity, demonstrating that replaced *Pa*LipA can be re-activated by *Pa*Lif.

6.2.4 *Pa*Lif binding affects the structural stability of key regions of *Pa*LipA involved in the opening of the active site

Our above shown results suggest that *Pa*Lif facilitates the opening of the active site in *Pa*LipA and stabilizes the partial β-sheet structure in the region of residues 17-30. To understand the

underlying mechanism how *Pa*Lif binding influences the active site opening in *Pa*LipA, we analyzed changes in the structural rigidity of *Pa*LipA upon mutating residues of *Pa*Lif that interact with *Pa*LipA using an ensemble- and rigidity-theory based perturbation approach [194] integrated into the CNA approach [211]. Initially, we identified interactions between *Pa*Lif and helix α5 as well as residues 17-30 of *Pa*LipA based on the Cα-Cα distance matrix averaged over the six unbiased MD simulations of the *Pa*Lif:LipA complex (Figure 23A). In total, 13 residues of *Pa*Lif were identified (195-203, 213, 217-220) that are in direct contact with the region of residues 17-30 of *Pa*LipA (Figure 23B). By contrast, no residues of *Pa*Lif were identified that interact with helix α5.

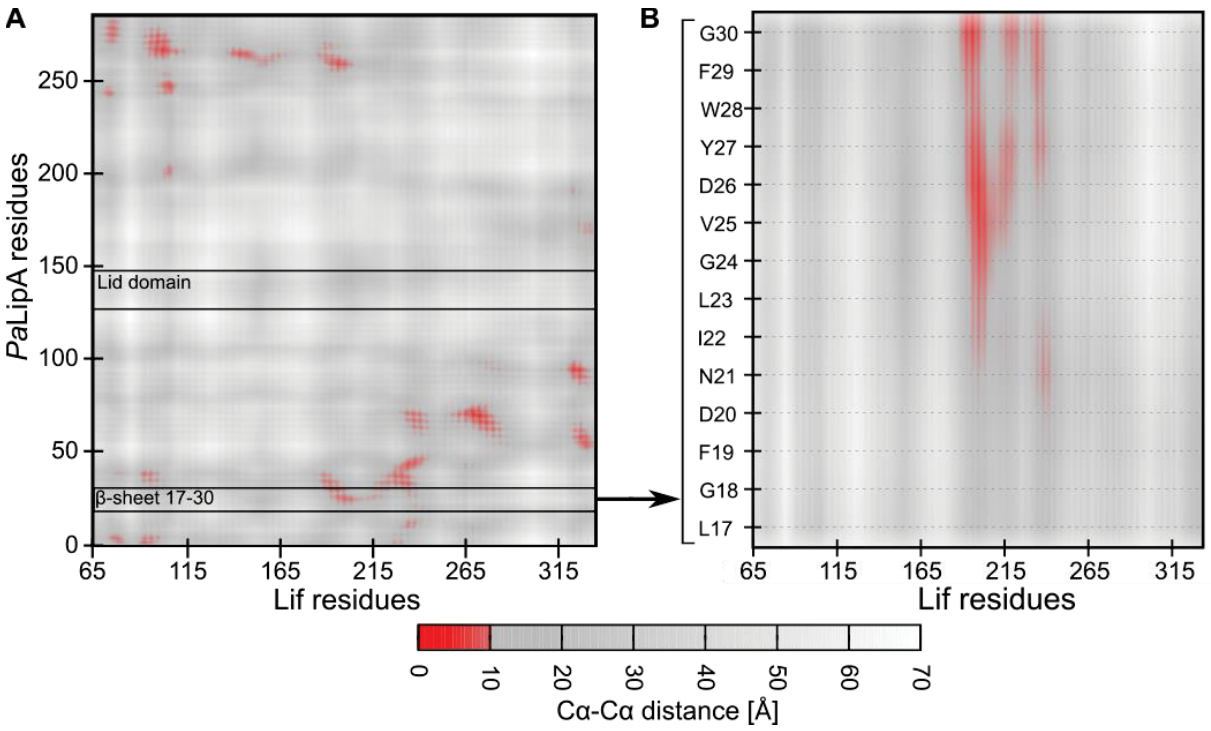


Figure 23. *Pa*Lif residues interacting with *Pa*LipA. (A) Average Cα-Cα distance matrix calculated for the *Pa*Lif:LipA complex over six unbiased MD simulations of 1 µs length each. Residue pairs with a Cα-Cα distance < 10 Å are colored in red (see color scale) and considered in direct contact. Regions of helix α5 and residues 17-30 in *Pa*LipA are indicated by black lines. The SEM is < 0.1 Å in all cases. **(B)** Close-up of the Cα-Cα distance matrix for residues 17-30 in *Pa*LipA. Color code as in panel A.

To probe a potential influence of *Pa*Lif binding on *Pa*LipA stability, first, a conformational ensemble of the *Pa*Lif:LipA complex was generated from the above shown MD simulations, constituting the ground state. A perturbed state of the *Pa*Lif:LipA complex was then generated

by removing the side chain of a *Pa*Lif residue except the Cβ atom, mimicking a substitution to alanine, but keeping the structures of *Pa*Lif and *Pa*LipA unchanged otherwise. This perturbation was carried out separately for each of the 13 above mentioned residues. The changes are quantified as residue-wise free energy $\Delta G_{i,\mathrm{CNA}}$ (**section 2.6.6**), a measure for structural stability [211]. Of the 13 residues tested, F195, R199, R203, D218, and R219 showed the largest effect on the structural stability of *Pa*LipA (Figure 24). Upon perturbation of residue F195, the changes in $\Delta G_{i,\mathrm{CNA}}$ were largest for *Pa*LipA residues 15-45, which form the oxyanion hole and the cleft of the active site, whereas residue R199 affects the stability of residues 15-45 and in addition residues 255-268, which constitute the loop stabilizing the catalytic triad residue H251 (Figures 24A and 24B). Upon perturbation of residue R203, in addition to residues 15- 45, residues 142-144, which constitute the neighboring loop at the C-terminus of helix α5 also showed substantial changes in $\Delta G_{i,\mathrm{CNA}}$ (Figure 24C). By contrast, residues D218 and R219 specifically affected the stability of region 17-30 of *Pa*LipA (Figures 24D and 24E). Notably, all affected residues belong to the substrate binding site of *Pa*LipA, which undergoes conformational rearrangements during activation. This finding indicates that the residues of *Pa*Lif that directly interact with *Pa*LipA lead to a long-range impact on the structural stability of *Pa*LipA regions (residues 142-144, 255-268 and 15-45) in the vicinity of *Pa*LipA's active site.

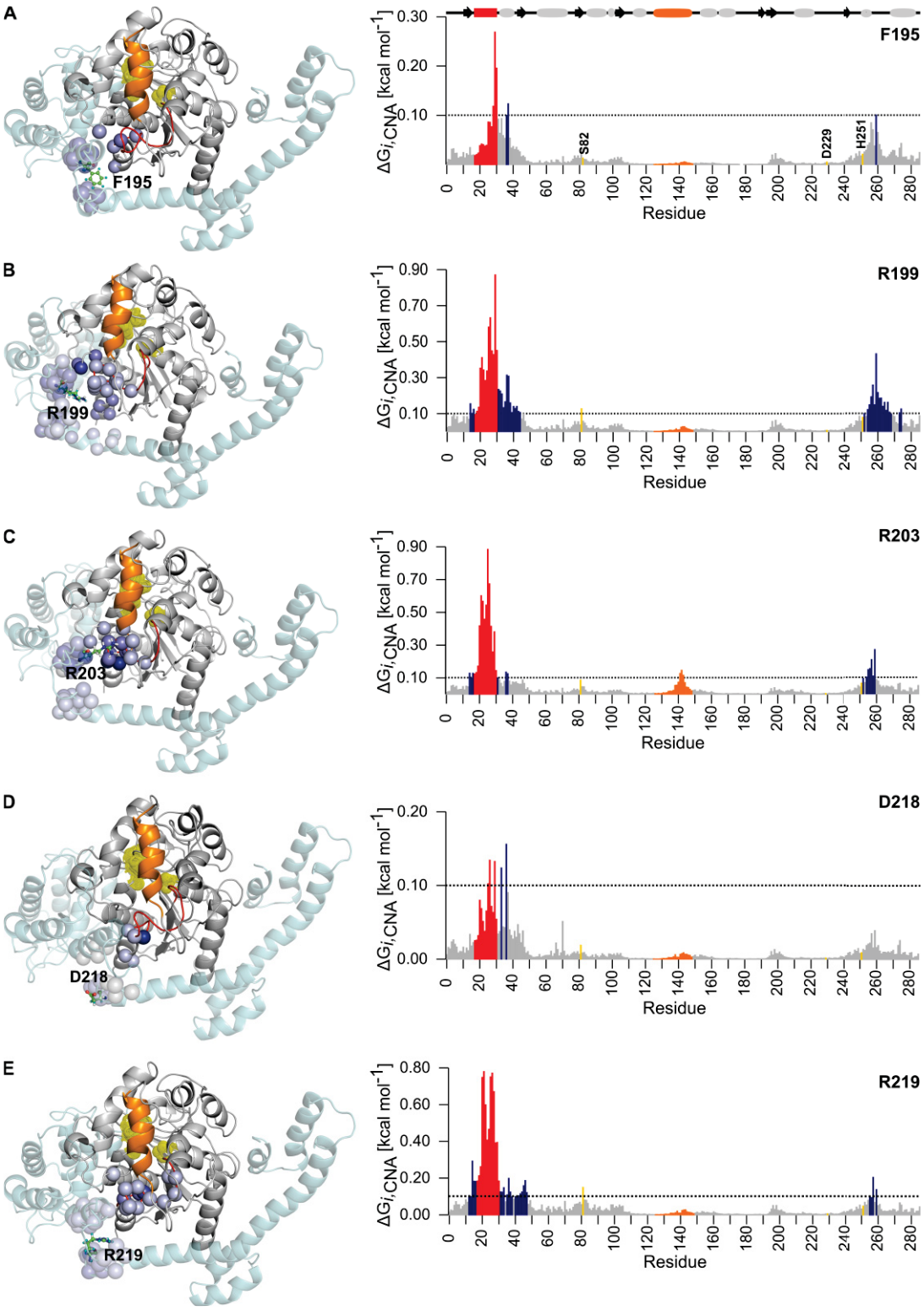


Figure 24. Potential influence of *Pa*Lif residues interacting with *Pa*LipA on the structural stability of *Pa*LipA. A perturbation approach implemented in CNA was applied on the ensemble of structures of *Pa*Lif:LipA generated by six unbiased MD simulations. **(A)** Left: Residues with $\Delta G_{i,\mathrm{CNA}}$ above the threshold of 0.1 kcal mol^{-1} are depicted as spheres on the *Pa*Lif:LipA complex structure. Blue colors reflect predicted $\Delta G_{i,\mathrm{CNA}}$ values, the larger the value, the darker is the color. The perturbed residue F195 of *Pa*Lif (green, ball and stick representation) influences the stability of residues 17-30 of *Pa*LipA (red). Helix α5 (orange) is shown in closed conformation occluding the binding site (yellow). Right: The histogram shows the per-residue $\Delta G_{i,\mathrm{CNA}}$ for *Pa*LipA. The dashed line indicates the threshold value of 0.1 kcal mol^{-1} above which residues are considered affected in terms of their

structural stability. Residues forming helix α5, the catalytic triad, and region 17-30 are highlighted in orange, yellow, and red, respectively. Other residues with $\Delta G_{i,CNA}$ above the threshold are highlighted in blue. **(B)** As in panel A for the perturbation of residue R199 of *Pa*Lif. **(C)** As in panel A for the perturbation of residue R203 of *Pa*Lif. **(D)** As in panel A for the perturbation of residue D218 of *Pa*Lif. **(E)** As in panel A for the perturbation of R219 of *Pa*Lif. The standard error of the mean is < 0.05 kcal mol^{-1} for all residues in all cases.

6.3 Conclusion and significance

In this study, we have shown at the atomistic level that the steric chaperone *Pa*Lif catalyzes the activation process of *Pa*LipA by structurally stabilizing the intermediate *Pa*LipA conformation, particularly the β-sheet in the region of residues 17-30, such that the opening of the lid domain is facilitated. This opening allows substrate access to *Pa*LipA's catalytic site.

The principal results of this study are:

- A partially open conformation was reached when starting MD simulations from either the closed or the open state of *Pa*LipA.
- The partially open state is more favored than the closed state of *Pa*LipA when bound to *Pa*Lif.
- The open state of *Pa*LipA is disfavored in terms of free energy compared to the closed state. However, in the *Pa*Lif:LipA complex, the open state is relatively more favorable than in free *Pa*LipA.
- *Pa*Lif significantly favors the formation of the β-sheet in the region of residues 17-30 in *Pa*LipA.
- The closed-to-open transition of *Pa*LipA is a reversible process, and *Pa*Lif is required for the conformational transition of *Pa*LipA to the open state as well as to stabilize *Pa*LipA in the open conformation under *in vitro* conditions.
- Specific *Pa*Lif residues that directly interact with *Pa*LipA lead to a long-range impact on the structural stability of certain regions (residues 142-144, 255-268 and 15-45) of *Pa*LipA in the vicinity of the active site. In particular, the partial β-sheet structure of residues 17-30 is affected, which has been shown to foster the opening of the active site.

Our results shed light on the molecular mechanism of a steric chaperone. They explain how *Pa*Lif directly catalyzes the folding process of *Pa*LipA by imprinting the essential steric (structural) information onto the target protein. *Pa*Lif structurally stabilizes an intermediate *Pa*LipA conformation, particularly a β-sheet in the region of residues 17-30, such that the opening of *Pa*LipA's lid domain is facilitated.

7. SUMMARY AND PERSPECTIVES

This thesis aimed to understand the molecular mechanism of *Pa*LipA activation mediated by its specific foldase *Pa*Lif. In order to achieve that, I applied MD simulations, CNA and free energy calculations in combination with experimental data to improve our understanding of the energetics and dynamics of the activation process.

***Pa*Lif- a structurally minimalistic, highly flexible steric chaperone**

Until now, Lif appears to be non-crystallizable in the absence of its cognate lipase. All attempts of many structural biology groups remained unsuccessful. In this work, first, I computationally investigated the structural dynamics of the individual domains of *Pa*Lif in the absence of *Pa*LipA to correlate the structural-functional dynamics (**section 4; publication I**). My findings revealed that MD1 of *Pa*Lif is structurally stable in both free as well as bound form. By contrast, MD2 and EHD both undergo helix-loop transitions in the free form and gain helicality upon binding. Overall, *Pa*Lif adopted a more compact state but still highly flexible conformations in the free form. Further, we reweighted our simulations with data from FRET experiments performed in the lab of Prof. Dr. C. Seidel. In doing so, a representative simulation-based structural ensemble of *Pa*Lif conformations was obtained, which increased the resolution of our FRET data to the atomistic level. These conformations showed that MD2 and EHD drive the compaction of this chaperone in free state. Finally, I computed the degree of burial of the interface and hydrophobic SASA and number of polar interactions formed, for the representative ensemble of free *Pa*Lif in comparison to *Pa*Lif:LipA complex. It was found that compaction of the structure is majorly driven by the unsatisfied polar interaction in free *Pa*Lif. The partially buried interface and hydrophobic SASA in compact conformations indicates that *Pa*Lif holds the affinity to re-open in order to bind *Pa*LipA despite the fickleness of the structure. However, it remains elusive whether binding of *Pa*LipA to *Pa*Lif is driven by an ‘induced fit’ mechanism or a consequence of ‘conformational selection’.

MD1 of *Pa*Lif- a binding platform for *Pa*LipA

As it was shown that MD1 is essential for the activation of *Pa*LipA [112], which was further corroborated by its structural stability (**section 4**), we investigated its role in the activation of *Pa*LipA and the underlying mechanism (**section 5; publication II**). Particularly, the *Pa*Lif$_{Y99A}$ variant was investigated, first, in terms of binding affinity by *in vitro* experiments done in the lab of Prof. Dr. K.-E. Jaeger and Prof. Dr. Claus Seidel. A stronger binding was observed in the case of *Pa*Lif (K_D = 39 ± 13 nM) as compared to the *Pa*Lif$_{Y99A}$ variant (K_D = 102 ± 35 nM). The changes in structural rigidity of *Pa*LipA upon Y99A mutation were computed using the CNA approach on the conformational ensemble generated from MD simulations for *Pa*Lif:LipA. This allowed us to identify a long-range network of interactions bridging from Y99 of *Pa*Lif to the cleft of the active site of *Pa*LipA formed by residues 17-30. Moreover, it indicates that a structural stabilization of this region might be essential for *Pa*LipA activation. Furthermore, solution NMR structures of MD1 and the MD1$_{Y99A}$ variant have been solved in the lab of Prof. Dr. Manuel Etzkorn, refined via ff14SB force field with NMR restraints and deposited in the PDB (PDB code: 5OVM; BMRB code: 34175 for MD1 and PDB code: 6GSF; BMRB code. 34286 for MD1$_{Y99A}$) (**section 5**; **publication II**).

***Pa*Lif facilitates *Pa*LipA activation by stabilizing the intermediate state**

Until now, our minimalistic knowledge about bacterial steric chaperones was restricted to the protease propeptides and fimbrial chaperones, which function intramolecularly [13]. Our research has added one more example of a membrane-bound steric chaperone, which executes its action by intermolecular interactions. Finally, to understand the underlying mechanism of *Pa*LipA activation via *Pa*Lif, first, I used unbiased MD simulations, which allowed me to correlate the structural changes and the functional dynamics of *Pa*LipA alone and in complex with *Pa*Lif (**section 6.2.1; publication III**). However, an understanding of the underlying energetics is also required. Therefore, a PMF was derived for the closed-to-open transition of

the lid in *Pa*LipA and the *Pa*Lif:LipA complex with biased MD simulations. A quantitative assessment of the active site opening in terms of free energies was obtained. My findings showed that the overall open state of *Pa*LipA is disfavored compared to the closed state. Yet, in the *Pa*Lif:LipA complex, the open state is more favorable than in *Pa*LipA alone (**section 6.2.2**). Furthermore, I also showed that the binding of *Pa*Lif significantly increases the formation of the partial β-sheet structure in the region of residues 17-30 of *Pa*LipA (**section 6.2.2**). These findings, in combination with the *in vitro* experiments done in the lab of Prof. Dr. K.-E. Jaeger, led us to conclude that the closed-to-open transition of *Pa*LipA is a reversible process. We further determined that *Pa*Lif is required for the conformational transition of *Pa*LipA to the open state and for stabilization of the open conformation until it is secreted into the extracellular medium (**section 6.2.3**). Lastly, an *in silico* perturbation analysis was performed on the directly interacting residues of *Pa*Lif and *Pa*LipA using the CNA approach (**section 6.2.4**). This analysis revealed that these residues lead to a long-range impact of the structural stability in the vicinity of *Pa*LipA's active site. In particular, these interactions stabilize the partial β-sheet structure formation of residues 17-30, which has been previously shown to foster the opening of the active site.

Altogether, my work provides valuable insights into the molecular mechanism of action of a steric chaperone. It provides a detailed explanation of how *Pa*Lif directly catalyzes the activation process of *Pa*LipA by imprinting the essential steric information, i.e., a partial β-sheet structure onto the target protein. In conclusion, *Pa*Lif structurally stabilizes the intermediate conformation of *Pa*LipA by stabilizing the structural changes required for the opening of the active site, which fosters the activation of *Pa*LipA.

_Pa_Lif:LipA- unravel kinetics and dynamics

One major question that has been raised and remains unanswered is how the Xcp-secreton is involved in *Pa*LipA secretion. In the future, insights into this process may provide valuable

information to unveil the remaining hidden kinetics of this system. A detailed understanding of the *Pa*Lif-mediated biogenesis of *Pa*LipA is of great importance for comprehensive perception of the complete molecular mechanism of the active *Pa*LipA production. Therefore, understanding the mode of action of these highly specific steric chaperones is of particular interest to be able to manipulate the efficiency and specificity of lipase secretion.

8. ACKNOWLEDGMENT

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Dr. Holger Gohlke for his continuous support, invaluable guidance, his patience, and motivation throughout my work. His guidance helped me in all the challenges and hard times of research, writing this thesis and my life in Germany. I could not have imagined having a better advisor, consultant, and mentor for my Ph.D. study.

As we say in India:- **"Gurur brahma, Gurur vishnu; Gurur devo mahesh varah**

Gurur sakshat param brahma; Tasmay shri gurve namah"

Besides my advisor, I would like to thank the rest of my project team, my collaborators Prof. Dr. K-E Jaeger, Prof. Dr. Claus Seidel, Prof. Dr. Manuel Etzkorn, Dr. Filip Kovacic, Dr. Jakub Kubaik, Dr. Peter Dollinger, and Dr. Aldino Viegas for their insightful discussions and hard questions.

I am heartily thankful to all of my colleagues and friends of the CPClab for providing awesome working atmosphere. Mainly I like to thank: Michele Bonus and Benedikt Frieg as my "second mentors", "Sambucca men" and "crazy carnival guys", which provided me a great start in Germany. Stephan Schott, Dr. Denis Schmidt, Mykola Dimura, Daniel Mulnaes, Dr. Christopher Pfleger and Anuseema for their valuable help in work done during Ph.D. and in this thesis. Jonas Dittrich and Bastian Schepers for reading this thesis and for very interesting discussions which brighten up my every day in the office and much more. To Tobias, Jagmohan, Giulia, Sabah, Nicola, Markus, David, Susanne and Christina for all the drinks we had together.

Pankaj Kumar for being my side in all ups and downs. Last but not least, my parents, my sister-in-law Shikha Verma and my brother Anurag Deepak Verma, for being there as my emotional support all time.

9. REFERENCES

1. Jaeger, K.-E. and Eggert, T., *Lipases for biotechnology.* Current Opinion in Biotechnology, 2002. **13**(4): p. 390-397.
2. Jaeger, K.-E. and Reetz, M.T., *Microbial lipases form versatile tools for biotechnology.* Trends in Biotechnology, 1998. **16**(9): p. 396-403.
3. Filloux, A., Bally, M.*, et al.*, *Protein secretion in gram-negative bacteria - transport across the outer-membrane involves common mechanisms in different bacteria.* EMBO Journal, 1990. **9**(13): p. 4323-4329.
4. Arpigny, J.L. and Jaeger, K.-E., *Bacterial lipolytic enzymes: classification and properties.* Biochemical Journal, 1999. **343**(1): p. 177-183.
5. Jaeger, K.-E., Schneidinger, B.*, et al.*, *Lipase of Pseudomonas aeruginosa: molecular biology and biotechnological application.* Molecular Biology of Pseudomonads, 1996: p. 319-330.
6. Lederberg, J., *Pseudomonas*, Encyclopedia of Microbiology, San Diego. p. 876-891.
7. Messaoudi, A., Belguith, H.*, et al.*, *Classification of EC 3.1. 1.3 bacterial true lipases using phylogenetic analysis.* African Journal of Biotechnology, 2010. **9**(48): p. 8243-8247.
8. Du Plessis, D.J., Nouwen, N.*, et al.*, *The sec translocase.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 2011. **1808**(3): p. 851-865.
9. Rosenau, F., Tommassen, J.*, et al.*, *Lipase-specific foldases.* ChemBioChem, 2004. **5**(2): p. 152-161.
10. Filloux, A., *The underlying mechanisms of type II protein secretion.* Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2004. **1694**(1-3): p. 163-179.
11. Ollis, D.L., Cheah, E.*, et al.*, *The α/β hydrolase fold.* Protein Engineering, Design and Selection, 1992. **5**(3): p. 197-211.
12. Jaeger, K.-E., Dijkstra, B.*, et al.*, *Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases.* Annual Reviews in Microbiology, 1999. **53**(1): p. 315-351.
13. Kris, P., *Structure/function relationship of the lipase-foldase complex of Burkholderia glumae*. 2008, Vrije Universiteit Brussel.
14. Frenken, L., Egmond, M.R.*, et al.*, *Cloning of the Pseudomonas glumae lipase gene and determination of the active site residues.* Applied and Environmental Microbiology, 1992. **58**(12): p. 3787-3791.
15. Lang, D., Hofmann, B.*, et al.*, *Crystal Structure of a Bacterial Lipase fromChromobacterium viscosumATCC 6918 Refined at 1.6 Å Resolution.* Journal of Molecular Biology, 1996. **259**(4): p. 704-717.
16. Khan, F.I., Lan, D.*, et al.*, *The lid domain in lipases: Structural and functional determinant of enzymatic properties.* Frontiers in Bioengineering and Biotechnology, 2017. **5**: p. 16.
17. Nardini, M., Lang, D.A.*, et al.*, *Crystal structure of Pseudomonas aeruginosa lipase in the open conformation: the prototype for family I. 1 of bacterial lipases.* Journal of Biological Chemistry, 2000.
18. Deveer, A.M., Dijkman, R.*, et al.*, *A monolayer and bulk study on the kinetic behavior of Pseudomonas glumae lipase using synthetic pseudoglycerides.* Biochemistry, 1991. **30**(41): p. 10034-10042.
19. Pauwels, K., Lustig, A.*, et al.*, *Structure of a membrane-based steric chaperone in complex with its lipase substrate.* Nature Structural and Molecular Biology, 2006. **13**(4): p. 374.

20. Pauwels, K., del Pino, M.M.S., *et al.*, *Decoding the folding of Burkholderia glumae lipase: folding intermediates en route to kinetic stability.* PloS One, 2012. **7**(5).
21. Rozycki, B., Kim, Y.C., *et al.*, *SAXS ensemble refinement of escrt-iii chmp3 conformational transitions.* Structure, 2011. **19**(1): p. 109-116.
22. Van Delden, C. and Iglewski, B.H., *Cell-to-cell signaling and Pseudomonas aeruginosa infections.* Emerging Infectious Diseases, 1998. **4**(4): p. 551-560.
23. Wu, W., Jin, Y., *et al.*, *Chapter 41 - Pseudomonas aeruginosa*, Molecular Medical Microbiology (Second Edition), 2015. Academic Press: Boston. p. 753-767.
24. Stover, C.K., Pham, X.Q., *et al.*, *Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.* Nature, 2000. **406**: p. 959.
25. Nouwens, A.S., Cordwell, S.J., *et al.*, *Complementing genomics with proteomics: the membrane subproteome of Pseudomonas aeruginosa PAO1.* Electrophoresis, 2000. **21**(17): p. 3797-3809.
26. Kung, V.L., Ozer, E.A., *et al.*, *The accessory genome of Pseudomonas aeruginosa.* Microbiology and Molecular Biology Reviews, 2010. **74**(4): p. 621-641.
27. Tsang, K.W., Shum, D.K., *et al.*, *Pseudomonas aeruginosa adherence to human basement membrane collagen in vitro.* European Respiratory Journal, 2003. **21**(6): p. 932-938.
28. Lamont, I.L., Beare, P.A., *et al.*, *Siderophore-mediated signaling regulates virulence factor production in Pseudomonas aeruginosa.* Proceedings of the National Academy of Sciences, 2002. **99**(10): p. 7072-7077.
29. Klockgether, J. and Tümmler, B., *Recent advances in understanding Pseudomonas aeruginosa as a pathogen.* F1000Research, 2017. **6**: p. 1261-1261.
30. Michalska, M. and Wolf, P., *Pseudomonas Exotoxin A: optimized by evolution for effective killing.* Frontiers in Microbiology, 2015. **6**(963).
31. Antonella, M., Anne, B.K., *et al.*, *Pseudomonas aeruginosa Pyocyanin Induces Neutrophil Death via Mitochondrial Reactive Oxygen Species and Mitochondrial Acid Sphingomyelinase.* Antioxidants & Redox Signaling, 2015. **22**(13): p. 1097-1110.
32. Barker, A.P., Vasil, A.I., *et al.*, *A novel extracellular phospholipase C of Pseudomonas aeruginosa is required for phospholipid chemotaxis.* Molecular Microbiology, 2004. **53**(4): p. 1089-1098.
33. Basso, P., Ragno, M., *et al.*, *Pseudomonas aeruginosa pore-forming exolysin and type IV pili cooperate to induce host cell lysis.* American Society of Microbiology, 2017. **8**(1): p. e02250-16.
34. Flaugnatti, N., Le, T.T.H., *et al.*, *A phospholipase A1 antibacterial Type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery.* Molecular Microbiology, 2016. **99**(6): p. 1099-1118.
35. Muthusamy, K., Gopalakrishnan, S., *et al.*, *Biosurfactants: properties, commercial production and application.* Current Science (00113891), 2008. **94**(6).
36. Haritash, A. and Kaushik, C., *Biodegradation aspects of polycyclic aromatic hydrocarbons (PAHs): a review.* Journal of Hazardous Materials, 2009. **169**(1-3): p. 1-15.
37. Thavasi, R., Subramanyam Nambaru, V.R.M., *et al.*, *Biosurfactant production by Pseudomonas aeruginosa from renewable resources.* Indian Journal of Microbiology, 2011. **51**(1): p. 30-36.
38. Costerton, J., Ingram, J., *et al.*, *Structure and function of the cell envelope of gram-negative bacteria.* Bacteriological Reviews, 1974. **38**(1): p. 87.
39. Silhavy, T.J., *The bacterial cell envelope.* Cold Spring Harbor Perspectives in Biology, 2010. **2**: p. a000414.

40. Lugtenberg, B. and Van Alphen, L., *Molecular architecture and functioning of the outer membrane of Escherichia coli and other gram-negative bacteria.* Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes, 1983. **737**(1): p. 51-115.
41. Douzi, B., Ball, G.*, et al.*, *Deciphering the Xcp Pseudomonas aeruginosa type II secretion machinery through multiple interactions with substrates.* Journal of Biological Chemistry, 2011. **286**(47): p. 40792-40801.
42. Miller, S.I. and Salama, N.R., *The gram-negative bacterial periplasm: Size matters.* PLoS Biology, 2018. **16**(1): p. e2004935.
43. Arends, S., Williams, K.*, et al.*, *The Periplasm*. 2007, ASM Press Washington, DC:.
44. Collins, R.F., Beis, K.*, et al.*, *The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in Escherichia coli.* Proceedings of the National Academy of Sciences, 2007. **104**(7): p. 2390-2395.
45. Marlovits, T.C., Kubori, T.*, et al.*, *Structural insights into the assembly of the type III secretion needle complex.* Science, 2004. **306**(5698): p. 1040-1042.
46. Henry, A.J., McDonnell, J.M.*, et al.*, *Conformation of the isolated Cε3 domain of IgE and its complex with the high-affinity receptor, FcεRI.* Biochemistry, 2000. **39**(25): p. 7406-7413.
47. Li, Z., Clarke, A.J.*, et al.*, *A major autolysin of Pseudomonas aeruginosa: subcellular distribution, potential role in cell growth and division and secretion in surface membrane vesicles.* Journal of Bacteriology, 1996. **178**(9): p. 2479-2488.
48. Beveridge, T.J., *Structures of gram-negative cell walls and their derived membrane vesicles.* Journal of Bacteriology, 1999. **181**(16): p. 4725-4733.
49. Braun, V., *Covalent lipoprotein from the outer membrane of Escherichia coli.* Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes, 1975. **415**(3): p. 335-377.
50. Raetz, C.R. and Whitfield, C., *Lipopolysaccharide endotoxins.* Annual Review of Biochemistry, 2002. **71**(1): p. 635-700.
51. Dalbey, R.E. and Kuhn, A., *Protein traffic in Gram-negative bacteria–how exported and secreted proteins find their way.* FEMS Microbiology Reviews, 2012. **36**(6): p. 1023-1045.
52. Filloux, A., *Protein secretion systems in Pseudomonas aeruginosa: an essay on diversity, evolution, and function.* Frontiers in Microbiology, 2011. **2**: p. 155.
53. Xie, K. and Dalbey, R.E., *Inserting proteins into the bacterial cytoplasmic membrane using the Sec and YidC translocases.* Nature Reviews Microbiology, 2008. **6**(3): p. 234.
54. Breyton, C., Haase, W.*, et al.*, *Three-dimensional structure of the bacterial protein-translocation complex SecYEG.* Nature, 2002. **418**(6898): p. 662.
55. Fekkes, P. and Driessen, A.J., *Protein targeting to the bacterial cytoplasmic membrane.* Microbiology and Molecular Biology Reviews, 1999. **63**(1): p. 161-173.
56. Dilks, K., Rose, R.W.*, et al.*, *Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey.* Journal of Bacteriology, 2003. **185**(4): p. 1478-1483.
57. Berks, B.C., Palmer, T.*, et al.*, *Protein targeting by the bacterial twin-arginine translocation (Tat) pathway.* Current Opinion in Microbiology, 2005. **8**(2): p. 174-181.
58. Koronakis, V., Sharff, A.*, et al.*, *Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export.* Nature, 2000. **405**(6789): p. 914.
59. He, S.Y., Nomura, K.*, et al.*, *Type III protein secretion mechanism in mammalian and plant pathogens.* Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2004. **1694**(1-3): p. 181-206.
60. Minamino, T. and Pugsley, A.P., *Measure for measure in the control of type III secretion hook and needle length.* Molecular Microbiology, 2005. **56**(2): p. 303-308.
61. Fronzes, R., Christie, P.J.*, et al.*, *The structural biology of type IV secretion systems.* Nature Reviews Microbiology, 2009. **7**(10): p. 703.

62. Filloux, A., Hachani, A.*, et al.*, *The bacterial type VI secretion machine: yet another player for protein transport across membranes.* Microbiology, 2008. **154**(6): p. 1570-1583.
63. Junker, M., Besingi, R.N.*, et al.*, *Vectorial transport and folding of an autotransporter virulence protein during outer membrane secretion.* Molecular Microbiology, 2009. **71**(5): p. 1323-1332.
64. Cianciotto, N.P., *Type II secretion: a protein secretion system for all seasons.* Trends in Microbiology, 2005. **13**(12): p. 581-588.
65. Filloux, A., Michel, G.*, et al.*, *GSP-dependent protein secretion in gram-negative bacteria: the Xcp system of Pseudomonas aeruginosa.* FEMS Microbiology Reviews, 1998. **22**(3): p. 177-198.
66. Ball, G., Durand, É.*, et al.*, *A novel type II secretion system in Pseudomonas aeruginosa.* Molecular Microbiology, 2002. **43**(2): p. 475-485.
67. Robert, V., Filloux, A.*, et al.*, *Subcomplexes from the Xcp secretion system of Pseudomonas aeruginosa.* FEMS Microbiology Letters, 2005. **252**(1): p. 43-50.
68. Fullner, K.J. and Mekalanos, J.J., *Genetic Characterization of a new type IV-A pilus gene cluster found in both classical and el tor biotypes ofVibrio cholerae.* Infection and Immunity, 1999. **67**(3): p. 1393-1404.
69. Sauvonnet, N., Vignon, G.*, et al.*, *Pilus formation and protein secretion by the same machinery in Escherichia coli.* The EMBO Journal, 2000. **19**(10): p. 2221-2228.
70. Zimmer, J., Nam, Y.*, et al.*, *Structure of a complex of the ATPase SecA and the protein-translocation channel.* Nature, 2008. **455**(7215): p. 936.
71. Bleves, S., Gérard-Vincent, M.*, et al.*, *Structure-function analysis of xcpp, a component involved in general secretory pathway-dependent protein secretion inPseudomonas aeruginosa.* Journal of Bacteriology, 1999. **181**(13): p. 4012-4019.
72. Korotkov, K.V., Sandkvist, M.*, et al.*, *The type II secretion system: biogenesis, molecular architecture and mechanism.* Nature Reviews Microbiology, 2012. **10**(5): p. 336.
73. Nunn, D.N. and Lory, S., *Cleavage, methylation, and localization of the Pseudomonas aeruginosa export proteins XcpT,-U,-V, and-W.* Journal of Bacteriology, 1993. **175**(14): p. 4375-4382.
74. Hobbs, M. and Mattick, J.S., *Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes.* Molecular Microbiology, 1993. **10**(2): p. 233-243.
75. Shevchik, V.E., Robert-Baudouy, J.*, et al.*, *Specific interaction between OutD, an Erwinia chrysanthemi outer membrane protein of the general secretory pathway, and secreted proteins.* The EMBO Journal, 1997. **16**(11): p. 3007-3016.
76. Macrae, A., *Lipase-catalyzed interesterification of oils and fats.* Journal of the American Oil Chemists' Society, 1983. **60**(2Part1): p. 291-294.
77. Sharma, B., Dangi, A.K.*, et al.*, *Contemporary enzyme based technologies for bioremediation: a review.* Journal of Environmental Management, 2018. **210**: p. 10-22.
78. Kirchner, G., Scollar, M.P.*, et al.*, *Resolution of racemic mixtures via lipase catalysis in organic solvents.* Journal of the American Chemical Society, 1985. **107**(24): p. 7072-7076.
79. Parida, S. and Dordick, J.S., *Substrate structure and solvent hydrophobicity control lipase catalysis and enantioselectivity in organic media.* Journal of the American Chemical Society, 1991. **113**(6): p. 2253-2259.
80. El Khattabi, M., Ockhuijsen, C.*, et al.*, *Specificity of the lipase-specific foldases of gram-negative bacteria and the role of the membrane anchor.* Molecular and General Genetics MGG, 1999. **261**(4-5): p. 770-776.

81. Gopinath, S.C.B., Anbu, P.*, et al.*, *Strategies to characterize fungal lipases for applications in medicine and dairy industry.* BioMed Research International, 2013. **2013**: p. 154549-154549.
82. Berg, J.M., Tymoczko, J.L.*, et al.*, *Biochemistry/Jeremy M. Berg, John L. Tymoczko, Lubert Stryer; with Gregory J. Gatto, Jr*. 2012, New York: WH Freeman.
83. Jaeger, K.E., Ransac, S.*, et al.*, *Bacterial lipases.* FEMS Microbiology Reviews, 1994. **15**(1): p. 29-63.
84. Angkawidjaja, C. and Kanaya, S., *Family I. 3 lipase: bacterial lipases secreted by the type I secretion system.* Cellular and Molecular Life Sciences CMLS, 2006. **63**(23): p. 2804-2817.
85. El Khattabi, M., Van Gelder, P.*, et al.*, *Role of the lipase-specific foldase of Burkholderia glumae as a steric chaperone.* Journal of Biological Chemistry, 2000. **275**(35): p. 26885-26891.
86. Kim, D.Y. and Dordick, J.S., *Combinatorial array-based enzymatic polyester synthesis.* Biotechnology and Bioengineering, 2001. **76**(3): p. 200-206.
87. Tan, T., Lu, J.*, et al.*, *Biodiesel production with immobilized lipase: a review.* Biotechnology Advances, 2010. **28**(5): p. 628-634.
88. Zhu, B. and Panek, J.S., *Methodology based on chiral silanes in the synthesis of polypropionate-derived natural products− total synthesis of Epothilone A.* European Journal of Organic Chemistry, 2001. **2001**(9): p. 1701-1714.
89. Athawale, V., Manjrekar, N.*, et al.*, *Enzymatic synthesis of chiral menthyl methacrylate monomer by Pseudomonas cepacia lipase catalysed resolution of (±)-menthol.* Journal of Molecular Catalysis B: Enzymatic, 2001. **16**(3-4): p. 169-173.
90. Thom, D., Swarthoff, T.*, et al.*, *Enzymatic detergent composition*. 1987, Google Patents.
91. Wolff, A. and Showell, M.S., *Application of lipases in detergents.* Surfactant Science Series, 1997: p. 93-106.
92. Jaeger, K.-E., *Protein technologies and commercial enzymes.* Current Opinion in Biotechnology, 2004. **4**(15): p. 269-271.
93. Schmid, A., Dordick, J.*, et al.*, *Industrial biocatalysis today and tomorrow.* Nature, 2001. **409**(6817): p. 258.
94. Houde, A., Kademi, A.*, et al.*, *Lipases and their industrial applications.* Applied Biochemistry and Biotechnology, 2004. **118**(1-3): p. 155-170.
95. Noble, M., Cleasby, A.*, et al.*, *The crystal structure of triacylglycerol lipase from Pseudomonas glumae reveals a partially redundant catalytic aspartate.* FEBS Letters, 1993. **331**(1-2): p. 123-128.
96. Lang, D., Stadler, P.*, et al.*, *Structural and kinetic investigations of enantiomeric binding mode of subclass i lipases from the family of Pseudomonadaceae.* 1999.
97. Lang, D.A., Mannesse, M.L.*, et al.*, *Structural basis of the chiral selectivity of Pseudomonas cepacia lipase.* European Journal of Biochemistry, 1998. **254**(2): p. 333-340.
98. Schrag, J.D., Li, Y.*, et al.*, *The open conformation of a Pseudomonas lipase.* Structure, 1997. **5**(2): p. 187-202.
99. Devescovi, G., Bigirimana, J.*, et al.*, *A clinical isolate of Burkholderia glumae causes severe disease symptoms in rice; involvement of a quorum sensing regulated secreted lipase.* Applied and Environmental Microbiology, 2007.
100. Rosenau, F. and Jaeger, K.-E., *Bacterial lipases from Pseudomonas: regulation of gene expression and mechanisms of secretion.* Biochimie, 2000. **82**(11): p. 1023-1032.
101. Pesci, E.C., Pearson, J.P.*, et al.*, *Regulation of las and rhl quorum sensing in Pseudomonas aeruginosa.* Journal of Bacteriology, 1997. **179**(10): p. 3127-3132.
102. Winson, M.K., Camara, M.*, et al.*, *Multiple N-acyl-L-homoserine lactone signal molecules regulate production of virulence determinants and secondary metabolites in*

Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences, 1995. **92**(20): p. 9427-9431.
103. Urban, A., Leipelt, M.*, et al.*, *DsbA and DsbC affect extracellular enzyme formation in Pseudomonas aeruginosa.* Journal of Bacteriology, 2001. **183**(2): p. 587-596.
104. Liebeton, K., Zacharias, A.*, et al.*, *Disulfide bond in Pseudomonas aeruginosalipase stabilizes the structure but is not required for interaction with its foldase.* Journal of Bacteriology, 2001. **183**(2): p. 597-603.
105. Schiene, C. and Fischer, G., *Enzymes that catalyse the restructuring of proteins.* Current Opinion in Structural Biology, 2000. **10**(1): p. 40-45.
106. Anfinsen, C.B., *Principles that govern the folding of protein chains.* Science, 1973. **181**(4096): p. 223-230.
107. Pauwels, K., Van Molle, I.*, et al.*, *Chaperoning Anfinsen: the steric foldases.* Molecular Microbiology, 2007. **64**(4): p. 917-922.
108. Bryan, P., Wang, L.*, et al.*, *Catalysis of a protein folding reaction: Mechanistic implications of the 2.0. ANG. structure of the subtilisin-prodomain complex.* Biochemistry, 1995. **34**(32): p. 10310-10318.
109. Sohl, J.L., Jaswal, S.S.*, et al.*, *Unfolded conformations of α-lytic protease are more stable than its native state.* Nature, 1998. **395**(6704): p. 817.
110. El Khattabi, M., Van Gelder, P.*, et al.*, *Role of the calcium ion and the disulfide bond in the Burkholderia glumae lipase.* Journal of Molecular Catalysis B: Enzymatic, 2003. **22**(5): p. 329-338.
111. Shibata, H. and Kato, H., *Molecular properties and activity of amino-terminal truncated forms of lipase activator protein.* Bioscience, Biotechnology, and Biochemistry, 1998. **62**(2): p. 354-357.
112. Radford, S.E., Laue, E.D.*, et al.*, *Conformational flexibility and folding of synthetic peptides representing an interdomain segment of polypeptide chain in the pyruvate dehydrogenase multienzyme complex of Escherichia coli.* Journal of Biological Chemistry, 1989. **264**(2): p. 767-775.
113. Shibata, H., Kato, H.*, et al.*, *Random mutagenesis on the Pseudomonas lipase activator protein, LipB: exploring amino acid residues required for its function.* Protein Engineering, 1998. **11**(6): p. 467-472.
114. Shibata, H., Kato, H.*, et al.*, *Calcium ion-dependent reactivation of a Pseudomonas lipase by its specific modulating protein, LipB.* The Journal of Biochemistry, 1998. **123**(1): p. 136-141.
115. Selkoe, D.J., *Folding proteins in fatal ways.* Nature, 2003. **426**(6968): p. 900.
116. Ross, C.A. and Poirier, M.A., *Protein aggregation and neurodegenerative disease.* Nature Medicine, 2004. **10**(7): p. S10.
117. Porebski, B.T., Keleher, S.*, et al.*, *Smoothing a rugged protein folding landscape by sequence-based redesign.* Scientific Reports, 2016. **6**: p. 33958.
118. Calderwood, S.K., Khaleque, M.A.*, et al.*, *Heat shock proteins in cancer: chaperones of tumorigenesis.* Trends in Biochemical Sciences, 2006. **31**(3): p. 164-172.
119. Feder, M.E. and Hofmann, G.E., *Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology.* Annual Review of Physiology, 1999. **61**(1): p. 243-282.
120. Saibil, H., *Chaperone machines for protein folding, unfolding and disaggregation.* Nature Reviews Molecular Cell Biology, 2013. **14**: p. 630.
121. John Ellis, R., *Molecular Chaperones: Opening and closing the Anfinsen cage.* Current Biology, 1994. **4**(7): p. 633-635.
122. Sauter, N.K., Mau, T.*, et al.*, *Structure of α-lytic protease complexed with its pro region.* Nature Structural & Molecular Biology, 1998. **5**(11): p. 945.

123. Nishiyama, M., Horst, R.*, et al.*, *Structural basis of chaperone–subunit complex recognition by the type 1 pilus assembly platform FimD.* The EMBO Journal, 2005. **24**(12): p. 2075-2086.
124. Kessler, E., Safrin, M.*, et al.*, *Elastase and the LasA Protease of Pseudomonas aeruginosa Are Secreted with Their Propeptides.* Journal of Biological Chemistry, 1998. **273**(46): p. 30225-30231.
125. Eder, J., Rheinnecker, M.*, et al.*, *Folding of subtilisin BPN': characterization of a folding intermediate.* Biochemistry, 1993. **32**(1): p. 18-26.
126. Sauer, F.G., Fütterer, K.*, et al.*, *Structural basis of chaperone function and pilus biogenesis.* Science, 1999. **285**(5430): p. 1058-1061.
127. Vetsch, M., Puorger, C.*, et al.*, *Pilus chaperones represent a new type of protein-folding catalyst.* Nature, 2004. **431**(7006): p. 329.
128. Jones, C.H., Danese, P.N.*, et al.*, *The chaperone-assisted membrane release and folding pathway is sensed by two signal transduction systems.* The EMBO Journal, 1997. **16**(21): p. 6394-6406.
129. Webb, B. and Sali, A., *Protein Structure Modeling with MODELLER*, in *Protein Structure Prediction*, Kihara, D., Editor. 2014, Springer New York: New York, NY. p. 1-15.
130. Eswar, N., Webb, B.*, et al.*, *Comparative Protein Structure Modeling Using Modeller.* Current Protocols in Bioinformatics, 2006. **15**(1): p. 5.6.1-5.6.30.
131. Mobarec, J.C., Sanchez, R.*, et al.*, *Modern homology modeling of g-protein coupled receptors: which structural template to use?* Journal of Medicinal Chemistry, 2009. **52**(16): p. 5207-5216.
132. Moult, J., *A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction.* Current Opinion in Structural Biology, 2005. **15**(3): p. 285-289.
133. Vlachakis, D., Bencurova, E.*, et al.*, *Chapter Seven - Current state-of-the-art molecular dynamics methods and applications*, Advances in Protein Chemistry and Structural, 2014, Academic Press. p. 269-313.
134. Karplus, M. and McCammon, J.A., *Molecular dynamics simulations of biomolecules.* Nature Structural Biology, 2002. **9**(9): p. 646-652.
135. Lesk, A.M., *Protein architecture: a practical approach*. Vol. 91. 1991: IRL Press Oxford:.
136. Kitao, A. and Go, N., *Investigating protein dynamics in collective coordinate space.* Current Opinion in Structural Biology, 1999. **9**(2): p. 164-169.
137. Salsbury, F.R., *Molecular dynamics simulations of protein dynamics and their relevance to drug discovery.* Current Opinion in Pharmacology, 2010. **10**(6): p. 738-744.
138. Adcock, S.A. and McCammon, J.A., *Molecular Dynamics: Survey of methods for simulating the activity of proteins.* Chemical Reviews, 2006. **106**(5): p. 1589-1615.
139. Durrant, J.D. and McCammon, J.A., *Molecular dynamics simulations and drug discovery.* BMC Biology, 2011. **9**(1): p. 71.
140. Robustelli, P., Piana, S.*, et al.*, *Developing a molecular dynamics force field for both folded and disordered protein states.* Proceedings of the National Academy of Sciences, 2018. **115**(21): p. E4758-E4766.
141. Shaw, D.E. *Anton: A specialized machine for millisecond-scale molecular dynamics simulations of proteins*. 2009 19th IEEE Symposium on Computer Arithmetic. 2009.
142. Ding, F., Tsao, D.*, et al.*, *Ab Initio Folding of Proteins with All-Atom Discrete Molecular Dynamics.* Structure, 2008. **16**(7): p. 1010-1018.
143. Bös, F. and Pleiss, J., *Multiple molecular dynamics simulations of TEM β-Lactamase: dynamics and water binding of the Ω-loop.* Biophysical Journal, 2009. **97**(9): p. 2550-2558.

144. Boomsma, W., Ferkinghoff-Borg, J.*, et al.*, *Combining experiments and simulations using the maximum entropy principle.* PLoS Computational Biology, 2014. **10**(2): p. e1003406.
145. Jaynes, E.T., *Information theory and statistical mechanics.* Physical Review, 1957. **106**(4): p. 620.
146. Shannon, C.E., *A mathematical theory of communication.* Bell System Technical Journal, 1948. **27**(3): p. 379-423.
147. Kullback, S. and RA, L., *On information and sufficiency.* The Annals of Mathematical Statistics, 1951. **22**: p. 79-86.
148. Thomas, M.U., *A generalized maximum entropy principle.* Operations Research, 1979. **27**(6): p. 1188-1196.
149. Gull, S.F. and Daniell, G.J., *Image reconstruction from incomplete and noisy data.* Nature, 1978. **272**(5655): p. 686.
150. Leung, H.T.A., Bignucolo, O.*, et al.*, *A rigorous and efficient method to reweight very large conformational ensembles using average experimental data and to determine their relative information content.* Journal of Chemical Theory and Computation, 2015. **12**(1): p. 383-394.
151. Dimura, M., Peulen, T.O.*, et al.*, *Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems.* Current Opinion in Structural Biology, 2016. **40**: p. 163-185.
152. Jares-Erijman, E.A. and Jovin, T.M., *FRET imaging.* Nature Biotechnology, 2003. **21**(11): p. 1387.
153. Roy, R., Hohng, S.*, et al.*, *A practical guide to single-molecule FRET.* Nature Methods, 2008. **5**(6): p. 507.
154. Harris, D.C., *Quantitative chemical analysis*. 2010: Macmillan.
155. Kalinin, S., Peulen, T.*, et al.*, *A toolkit and benchmark study for FRET-restrained high-precision structural modeling.* Nature Methods, 2012. **9**(12): p. 1218-U129.
156. Sindbert, S., Kalinin, S.*, et al.*, *Accurate distance determination of nucleic acids via forster resonance energy transfer: implications of dye linker length and rigidity.* Journal of the American Chemical Society, 2011. **133**(8): p. 2463-2480.
157. Peulen, T.O., Opanasyuk, O.*, et al.*, *Combining graphical and analytical methods with molecular simulations to analyze time-resolved fret measurements of labeled macromolecules accurately.* Journal of Physical Chemistry B, 2017. **121**(35): p. 8211-8241.
158. Kudryavtsev, V., Sikor, M.*, et al.*, *Combining MFD and PIE for accurate single-pair forster resonance energy transfer measurements.* Chemphyschem, 2012. **13**(4): p. 1060-1078.
159. Sisamakis, E., Valeri, A.*, et al.*, *Accurate single-molecule fret studies using multiparameter fluorescence detection.* Methods in Enzymology, Vol 475: Single Molecule Tools, Pt B, 2010. **475**: p. 455-514.
160. Gopich, I.V. and Szabo, A., *Theory of the energy transfer efficiency and fluorescence lifetime distribution in single-molecule FRET.* Proceedings of the National Academy of Sciences, 2012. **109**(20): p. 7747-7752.
161. Kalinin, S., Felekyan, S.*, et al.*, *Probability distribution analysis of single-molecule fluorescence anisotropy and resonance energy transfer.* Journal of Physical Chemistry B, 2007. **111**(34): p. 10253-10262.
162. Kilic, S., Felekyan, S.*, et al.*, *Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1 alpha.* Nature Communications, 2018. **9**.
163. Felekyan, S., Sanabria, H.*, et al.*, *Analyzing forster resonance energy transfer with fluctuation algorithms.* Fluorescence Fluctuation Spectroscopy, Pt B, 2013. **519**: p. 39-85.

164. Henzler-Wildman, K. and Kern, D., *Dynamic personalities of proteins.* Nature, 2007. **450**: p. 964.
165. Hooft, R.W.W., Eijck, B.P.v.*, et al.*, *An adaptive umbrella sampling procedure in conformational analysis using molecular dynamics and its application to glycol.* The Journal of Chemical Physics, 1992. **97**(9): p. 6690-6694.
166. Patey, G.N. and Valleau, J.P., *A Monte Carlo method for obtaining the interionic potential of mean force in ionic solution.* The Journal of Chemical Physics, 1975. **63**(6): p. 2334-2339.
167. Kästner, J., *Umbrella sampling.* Wiley Interdisciplinary Reviews: Computational Molecular Science, 2011. **1**(6): p. 932-942.
168. Roux, B., *The calculation of the potential of mean force using computer simulations.* Computer Physics Communications, 1995. **91**(1): p. 275-282.
169. van Gunsteren, W.F., Daura, X.*, et al.*, *Computation of free energy.* Helvetica Chimica Acta, 2002. **85**(10): p. 3113-3129.
170. Mülders, T., Krüger, P.*, et al.*, *Free energy as the potential of mean constraint force.* The Journal of Chemical Physics, 1996. **104**(12): p. 4869-4870.
171. Kumar, S., Rosenberg, J.M.*, et al.*, *THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method.* Journal of Computational Chemistry, 1992. **13**(8): p. 1011-1021.
172. Eggeling, C., Berger, S.*, et al.*, *Data registration and selective single-molecule analysis using multi-parameter fluorescence detection.* Journal of Biotechnology, 2001. **86**(3): p. 163-80.
173. Radestock, S. and Gohlke, H., *Constraint network analysis: exploiting the link between protein rigidity and thermostability.* Engineering in Life Sciences, 2008. **8**: p. 507-522.
174. Radestock, S. and Gohlke, H., *Protein rigidity and thermophilic adaptation.* Proteins: Structure, Function, and Bioinformatics, 2011. **79**(4): p. 1089-1108.
175. Pfleger, C., Rathi, P.C.*, et al.*, *Constraint Network Analysis (CNA): a Python software package for efficiently linking biomacromolecular structure, flexibility,(thermo-) stability, and function*. 2013, ACS Publications.
176. Jacobs, D.J., Rader, A.J.*, et al.*, *Protein flexibility predictions using graph theory.* Proteins: Structure, Function, and Bioinformatics, 2001. **44**(2): p. 150-165.
177. Rader, A., Hespenheide, B.M.*, et al.*, *Protein unfolding: rigidity lost.* Proceedings of the National Academy of Sciences, 2002. **99**(6): p. 3540-3545.
178. Thorpe, M., Lei, M.*, et al.*, *Protein flexibility and dynamics using constraint theory.* Journal of Molecular Graphics and Modelling, 2001. **19**(1): p. 60-69.
179. Jacobs, D.J. and Thorpe, M.F., *Generic rigidity percolation: the pebble game.* Physical Review Letters, 1995. **75**(22): p. 4051.
180. Hespenheide, B.M., Rader, A.*, et al.*, *Identifying protein folding cores from the evolution of flexible regions during unfolding.* Journal of Molecular Graphics and Modelling, 2002. **21**(3): p. 195-207.
181. Radestock, S. and Gohlke, H., *Exploiting the link between protein rigidity and thermostability for data-driven protein engineering.* Engineering in Life Sciences, 2008. **8**(5): p. 507-522.
182. Dahiyat, B.I., Benjamin Gordon, D.*, et al.*, *Automated design of the surface positions of protein helices.* Protein Science, 1997. **6**(6): p. 1333-1337.
183. Pfleger, C., Minges, A.*, et al.*, *Ensemble-and rigidity theory-based perturbation approach to analyze dynamic allostery.* Journal of Chemical Theory and Computation, 2017. **13**(12): p. 6343-6357.
184. Rathi, P.C., Jaeger, K.-E.*, et al.*, *Structural rigidity and protein thermostability in variants of lipase A from Bacillus subtilis.* PloS One, 2015. **10**(7): p. e0130289.

185. Milić, D., Dick, M.*, et al.*, *Recognition motif and mechanism of ripening inhibitory peptides in plant hormone receptor ETR1.* Scientific Reports, 2018. **8**(1): p. 3890.
186. Jaeger, K., *Extracellular enzymes of Pseudomonas aeruginosa as virulence factors.* Immunitat und Infektion, 1994. **22**(5): p. 177-180.
187. Moglich, A., Joder, K.*, et al.*, *End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(33): p. 12394-12399.
188. Nettels, D., Hoffmann, A.*, et al.*, *Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds.* Journal of Physical Chemistry B, 2012. **116**(19): p. 5801-5801.
189. Yirdaw, R.B. and Mchaourab, H.S., *Direct observation of t4 lysozyme hinge-bending motion by fluorescence correlation spectroscopy.* Biophysical Journal, 2012. **103**(7): p. 1525-1536.
190. Chung, H.S., McHale, K.*, et al.*, *Single-molecule fluorescence experiments determine protein folding transition path times.* Science, 2012. **335**(6071): p. 981-984.
191. Felekyan, S., Kalinin, S.*, et al.*, *Filtered FCS: Species auto- and cross-correlation functions highlight binding and dynamics in biomolecules.* Chemphyschem, 2012. **13**(4): p. 1036-1053.
192. Antonik, M., Felekyan, S.*, et al.*, *Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis.* Journal of Physical Chemistry B, 2006. **110**(13): p. 6970-6978.
193. Hornak, V., Abel, R.*, et al.*, *Comparison of multiple Amber force fields and development of improved protein backbone parameters.* Proteins: Structure, Function, and Bioinformatics, 2006. **65**(3): p. 712-725.
194. Hermans, S.M., Pfleger, C.*, et al.*, *Rigidity theory for biomolecules: concepts, software, and applications.* Wiley Interdisciplinary Reviews: Computational Molecular Science, 2017. **7**(4): p. e1311.
195. Lang, D., Hofmann, B.*, et al.*, *Crystal structure of a bacterial lipase from Chromobacterium viscosum ATCC 6918 refined at 1.6 angstroms resolution.* Journal of Molecular Biology, 1996. **259**(4): p. 704-17.
196. Trodler, P., Schmid, R.D.*, et al.*, *Modeling of solvent-dependent conformational transitions in Burkholderia cepacia lipase.* BMC Structural Biology, 2009. **9**(1): p. 38.
197. Lee, H.S., Oh, Y.*, et al.*, *Molecular basis of aqueous-like activity of lipase treated with glucose-headed surfactant in organic solvent.* The Journal of Physical Chemistry B, 2018. **122**(47): p. 10659-10668.

10. ORIGINAL PUBLICATION I (In preparation)

Functional dynamics of a structurally minimalistic chaperone

Jakub Kubiak[1#], Neha Verma[2#], Holger Gohlke[2,4,5], Claus Seidel[1], Peter Dollinger[3], Filip Kovacic[3], Karl-Erich Jaeger[3,6]

[1]Institute of Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; [2]Heinrich-Heine-University of Düsseldorf, Institute of Pharmaceutical and Medicinal Chemistry, Düsseldorf, Germany; [3]Institute for Complex Systems - Structural Biochemistry (ICS-6), Forschungszentrum Jülich GmbH, Jülich, Germany; [4]John von Neumann Institute for Computing (NIC), Jülich Supercomputing Center (JSC); [5]Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany; [6]Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

[#] These authors contributed equally to this work

Keywords*: *Pseudomonas aeruginosa*, molecular chaperone, steric chaperone, lipase, molecular dynamics simulations, FRET, protein folding, structural ensemble

Abstract

LipA requires a cytoplasmic membrane-bound steric chaperone called lipase-specific foldase (Lif) for the activation. Lif binds LipA inside the cell periplasm, acts as a conformational switch during LipA activation followed by the release and secretion of LipA across the outer membrane indicating notable dynamics upon the formation of the Lif:LipA complex. In the Lif:LipA complex, Lif forms a flexible α-helical scaffold embracing LipA in a headphone like structure. Here we show, that in the free form, Lif does not stay in the hollow ''headphone'' conformation but rather exhibits large-scale conformational dynamics, in which the α-helical structure undergoes reversible compactions and extensions on the microsecond to sub-microsecond timescale. We also show that in the free conformation, Lif loses its secondary structure. The details of the process are studied by fluorescence spectroscopic in combination with all-atom molecular dynamics simulations. To study the atomistic details of the free Lif, we combined the data from FRET and MD simulations and generated an ensemble of 35 structures by using the maximum entropy refinement method. The simulation-based ensemble refinement allowed us to find limiting states of free Lif that range from partially compact to completely compact states, based on the radius of gyration. The hydrophobic surface provides the basis for the reopening of free Lif. This study provides structural insights into the action of Lif as a highly specific steric chaperone.

Introduction

Proper folding and targeting within cells are needed for all organisms. Proteins are synthesized as an unfolded chain of amino acids and adopt a unique stable fold after visiting a huge number of possible conformations. Some proteins can fold reversibly to their active native state, whereas other proteins need helper protein chaperones to achieve their proper active fold. Classical chaperones, such as GroEL and Hsp90 or Hsp70, bind to unfolded or misfolded state

of protein which is minimally structured and guide the folding process by preventing or reversing incorrect interactions using ATP hydrolysis (figure 1A). [1, 2]

However, several bacterial extracellular lipases are synthesized as folded but inactive precursors and achieve their enzymatically active conformation only upon interaction with specific foldase proteins. [3, 4] These lipase foldase proteins (Lifs) are classified as steric chaperones as they directly catalyze the conformational transition to the active state in an ATP-independent manner instead of preventing the protein from incorrect folding (figure 1B). [5] These steric chaperones bind to the folded intermediate state of the protein and lower the high energy barrier of the state transition, hence increase the rate of activation. [2, 6] The activation of the lipase by its foldase is a reversible process, at least in *in vitro* conditions. [7] Therefore the secretion of the lipase by type II secretion system (Xcp complex) is an important step in this process. [4, 8].

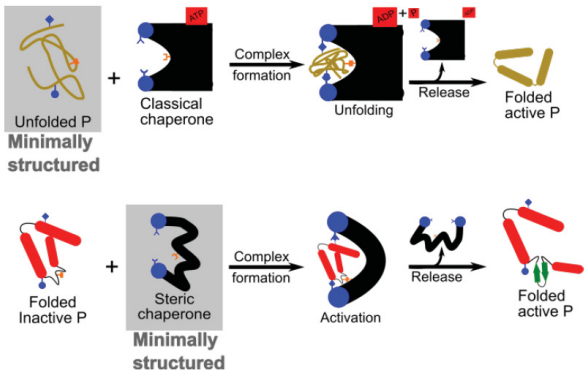


Figure 1. Molecular chaperones are key molecules in protein folding. (A) Classical chaperones bind to minimally structured target proteins and assist the folding by preventing off-pathway interactions. Generally, classical chaperones use ATP hydrolysis to reverse the folding process and do not provide any steric information required for correct folding. **(B)** Pre-folded inactive proteins need to bind minimalistically structured steric chaperones, and both undego required conformational changes. Hence, steric chaperones directly catalyzed the folding pathway.

Lipase A (LipA) produced by the gram-negative, pathogenic bacterium *Pseudomonas aeruginosa* is a virulence factor which not only affects plants and animals but also caused serious infections in humans. It releases inflammatory mediators from mouse peritoneal macrophages [9], provides nutrients by degrading lung surfactant lipids *in vivo* [10, 11], and

indirectly affects the production of pyoverdine, an important virulence factor. [12] To achieve its active native fold, LipA depends on an inner membrane integrated steric chaperone, the lipase-specific foldase (Lif). [13] However, how Lif assists the LipA folding is still an open question, and a comprehensive description of its structural preferences is a prerequisite for understanding the lipase folding.

Despite low sequence identity among lipase foldases (only eight amino acids are highly conserved among all known Lifs), a sequence-based structure prediction reveals that all Lifs possess similar structural folds, consisting of 70 % α-helices and 30 % random coils. [4]

Lifs are believed to have a common domain organization: a folding domain (FOD) responsible for interaction with the specific lipase in the periplasmic space of bacteria is connected by a flexible alanine- and proline-rich linker to a highly conserved transmembrane helix, which anchors them in the cytoplasmic membrane (figure 2A).[14, 15] *In vitro* experiments with truncated Lifs showed that only the FOD is needed for lipase activation.[16, 17] However, *in vivo*, the linker and transmembrane domain are needed to position Lif in the periplasmic space and for interactions with secretion machinery to facilitate secretion of LipA.[16, 18]

Pauwels et al. [19] solved the crystal structure of *Burkholderia glumae* lipase foldase (*Bg*Lif) folding domain (FOD) in complex with its cognate lipase (*Bg*LipA). *Bg*Lif has 39% identity and 52% similarity with Lif from *P. aeruginosa*. In the complexed state, FOD domain consists of 11 α-helices connected with loops and embracing *Bg*LipA in a headphone like structure with two globular mini-domains (MD1 and MD2) at the N- and C-terminal ends. Each mini-domain consists of three α-helices (MD1: H1-H3, MD2: H9-H11) connected by an extended α-helical domain (EHD) (H4-H8) (figure 2A).

Far- and near-UV CD spectroscopy indicated that *Bg*Lif undergoes large structural changes upon interaction with *Bg*LipA, which are accompanied by the formation of α-helical structures

presumably in the EHD. [2] This all indicates the dynamical role of Lif in the activation of lipase.

We investigate the conformational dynamics that lipase foldase from *P. aeruginosa* (Lif) undergoes in the free state compared to LipA-bound state by combined high-precision Förster resonance energy transfer (hpFRET) spectroscopy [20] and all-atom molecular dynamics (MD) simulations. Our results showed that in the free state, Lif partially preserves elements of secondary structure but lacks a stable tertiary structure. While MD1 remains relatively stable, MD2 and EHD are minimally structured with Lif adopting several compact and extended conformations that remain in dynamic equilibrium. Our results shed light on the fundamental mechanism of action of lipase foldase and possibly the other related steric chaperones.

Results

Secondary structure content of isolated Lif domains, free and bound Lif

Based on previous CD spectroscopy experiments and secondary structure analyses of the *Bg*Lif:LipA complexes [2], it was postulated that binding of LipA leads to structural changes predominantly in the extended α-helical domain (EHD), while the mini-domains (MD1 and MD2) noticeably retain their structure [2, 19]. Here we aimed to determine the secondary structure content of the catalytic folding domain (FOD) in the absence of LipA via MD simulations to better understand the dynamics of free Lif. We used all-atom molecular dynamics (MD) simulations to study the structural stability of the isolated, free mini-domains (MD1 and MD2) and the spacer region (EHD) of Lif in explicit water (see Materials and Methods for details) (figure 2A). To account for a potential force field bias, the simulations were performed with the ff99SB [21] and ff14SB [22] force fields. With either force field, three independent MD simulations of 1 μs length were performed for each domain extracted from the structure of Lif (figure 2A) obtained by homology modeling, using *Bg*Lif as a template.

In the course of the MD simulations, MD1 retained almost completely the α-helical secondary structure (Table 1). The average α-helical content (AHC) of MD1 was respectively ~72 ± 1.0% (mean ± SEM) and ~84 ± 3.0% for ff99SB and ff14SB,), which is in good agreement with the secondary structure content of a recent NMR structure of free MD1 (PDB code: 5OVM). Both EHD and MD2 show a less persistent α-helical secondary structure in the course of the simulations (figure 2B), as reflected in the AHC of these domains of 45.0 ± 5% (76.4 ± 1%) and 55.9 ± 9% (70.4 ± 7%) for ff99SB (ff14SB) (Table 1). Some parts of EHD (especially H6, H7, and H8) and MD2 (especially H9 and H11) showed pronounced and repeated changes in the per-residue α-helix propensity during the MD simulations irrespective of the used force field (figure S12B), suggesting that in free Lif, EHD and MD2 undergo helix-coil transitions. The maximum difference in AHC between both force fields is found for α8 (Table 1); there, ff14SB results in a ~3 fold higher AHC than ff99SB, which is equivalent to a change in free energy of ~1 kT at the simulation temperature. Hence, it is still within the uncertainty of the force fields. [21] We also computed the difference in the average α-helix propensity per-residue of full-length Lif (omitting residues 1-65) in complex with LipA with respect to free Lif (figure 2C), comparing ten MD simulations of 1 μs length each for either state (all simulations were performed with the ff99SB force field). A significant increase in α-helix propensity up to 54% was seen in MD2 (H9) when Lif binds to LipA, whereas in the EHD both increases and decreases of AHC are found. These results suggest that MD2, and to some extent EHD, undergo conformational changes, and MD2 in particular gains secondary structure, upon binding to LipA.

Table 1 Averaged α-helix content (AHC) of free isolated domains of Lif.[a]

	AHC[a] [23]	
α-helix	**ff99SB**	**ff14SB**
H 1	73.9 ± 2	84.0 ± 3
H2	75.2 ± 2	84.2 ± 3
H3	66.3 ± 0	83.6 ± 1
MD1	71.8 ± 1	84.0 ± 3
H4	75.5 ± 3	92.9 ± 1
H5	50.3 ± 0	66.1 ± 2
H6	23.2 ± 0	53.7 ± 2
H7	50.0 ± 2	85.2 ± 1
H8	26.4 ± 3	84.1 ± 4
EHD	45.0 ± 5	76.4 ± 1
H9	66.3 ± 4	77.7 ± 2
H10	76.3 ± 3	98.9 ± 1
H11	25.1 ± 1	34.8 ± 1
MD2	55.9 ± 9	70.4 ± 7

[a] The AHC per-residue calculated over three independent MD simulations is given as mean ± SEM. The percentage of AHC is given for individual helices as well as for domains as mean ± SEM.

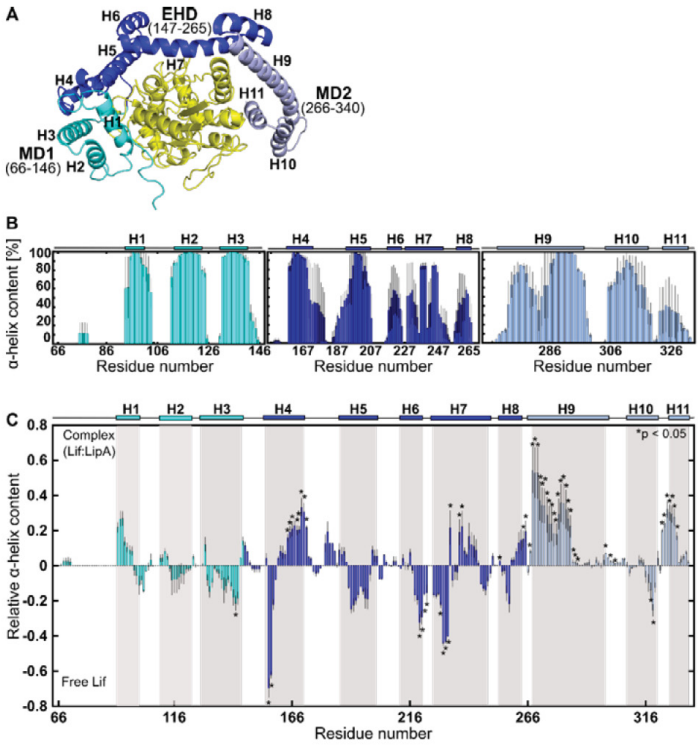


Figure 2. **Secondary structure analysis of Lif by MD simulations. (A)** Tertiary structure of Lif:LipA complex homology model showing how mini-domain 1 (MD1, residues 66-146) (cyan), mini-domain 2 (MD2, residues 147-265) (light blue), and extended helical domain (EHD, residues 266-340)) (blue) are defined in this article. The

α-helices of Lif are labeled H1 to H11, forming the novel fold around LipA (yellow). **(B)** α-Helical content per-residue averaged over three independent MD simulations for each of the three isolated domains MD1, EHD, and MD2, applying ff99SB. **(C)** The difference in the average per-residue AHC of Lif in complex with LipA with respect to free Lif computed over ten independent MD simulations. In **(B)** and **(C)**, error bars indicate SEM. *p < 0.05 calculated by two-sided t-test.

Dynamics of the super-tertiary structure of free Lif

MD simulations of the isolated domains of Lif indicated that mainly EHD and MD2 undergo conformational rearrangements in terms of secondary structure as compared to MD1. Next, to study the relative movements of these three domains, we evaluated the structural dynamics of FOD of free Lif from the ten independent all-atom MD simulations of 1 μs length each. The dihedral angle defined by four residues in MD1 (Q137), EHD (A215, E268), and MD2 (R296) (figure 3A), as a measure for a twisting motion of Lif, and the radius of gyration (R_G), as a measure of Lif's compactness, were computed. Across the ten MD simulations, Lif adopts dihedral angles covering the entire range from 0° to 360° (figure 3B), which indicates pronounced twisting motions. The time traces showed, however, only limited overlap of the sampled values among trajectories, which indicates that the twisting motions of free Lif are slower than the timescale covered by the MD simulations. In all MD simulations, a compaction of the structure to an R_G of ~21 Å was observed (figure 3C), starting from an R_G of 31 Å computed for free Lif in the bound conformation (figure 3A). Apparently, the compaction occurs on a time scale faster than a μs. Free Lif likely can populate multiple conformations ranging from an extended to a compact state with different dihedral angles (figure 3D).

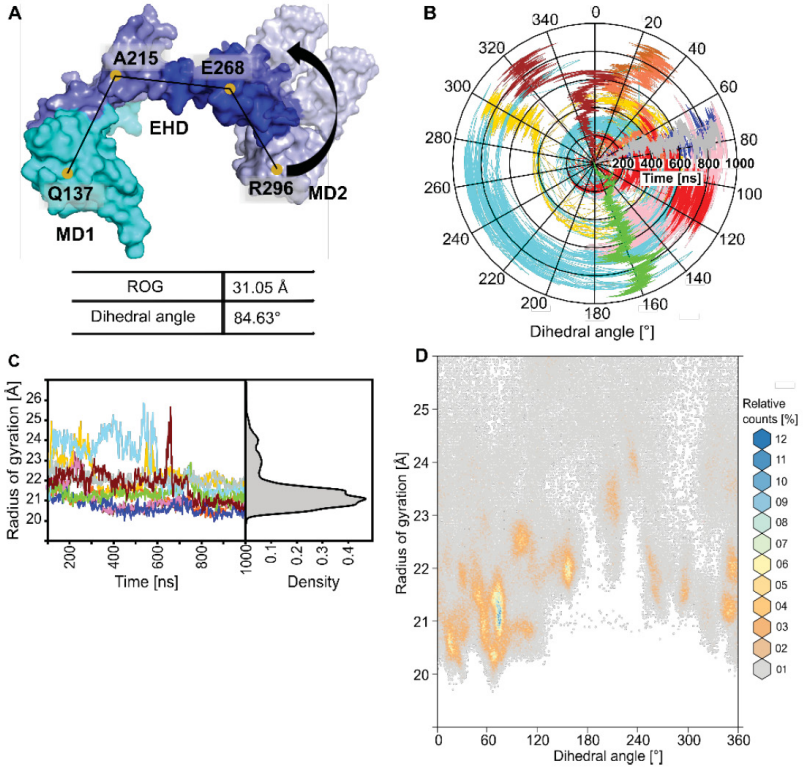


Figure 3. Structural dynamics of free Lif. (A) Location of the residues (Q137, A215, E268, and R296) forming the dihedral angle that was used as a reaction coordinate to describe the twisting motions of free Lif indicated by the black arrow. Domain coloring as in figure 2A. The dihedral angle and radius of gyration (R_G) for bound Lif are indicated below the structure. (B) Pronounced twisting motion (defined in panel A) of free Lif is notable from variation of dihedral angles as a function of simulation time (0 ns at the center, 1 µs at the rim). Each of ten independent MD simulations is indicated by different color. (C) R_G of free Lif as a function of time for ten independent MD simulations (different colors), of which the first 100 ns were discarded. On the right, the cumulative histogram is plotted. In all MD simulations, compaction of the structure was observed to an R_G of ~21 Å. (D) Distribution of conformations of free Lif generated by MD simulations mapped to the R_G / dihedral angle plane. Clusters of highly abundant conformations (orange) among a vast number of low abundant conformations (gray) indicate the complexity of the structural dynamics of free Lif.

Activity and affinity of fluorescently labeled Lif variants to LipA

To quantify the effects of Cys mutations and covalently bound fluorescent dye on the function of Lif we tested an *in vitro* efficiency of the six labeled double Cys Lif variants to convert inactive LipA into the catalytically active form using the lipase activity assay. Labeled double Cys variants of Lif revealed a comparable activation efficiency constant (K_{act}) for all six used Lif variants (Figures 3A and S5). Furthermore, labeled Lif variants were tested for their ability

to form an equilibrium complex with LipA based on the difference between average FRET efficiency of free Lif in solution and complex (Figure S6). The obtained K_D values for six Lif variants are more uniform than measured activation efficiency constants (figure 4B). It should be noted that although Cys substitution mutations and labeling do not affect the affinity of Lif to LipA, there is a detectable difference in the ability to activate LipA by different Lif variants as reported by K_{act}.

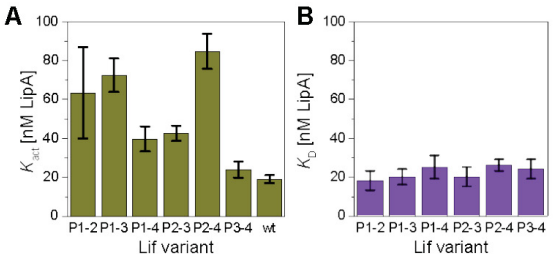


Figure 4. Double Cys variants of Lif labeled with Alexa488/647 activate and bind to LipA. (A) The activation constant (K_{act}), refereed as the effective concentration of LipA activated with labeled Lif variants, was determined by titrating labeled Lif variants with pre-active LipA (Figure S5). The results are mean ± SEM of three independent experiments, with each set in triplicates. (**B**) Equilibrium dissociation constant K_D of the complex investigated with FRET assay with labeled Lif variants. LipA binding is examined by monitoring the FRET signal from labeled Lif (Figure S6). Uncertainty is given by a standard error of the mean from the joined fitting of two experiments.

Fluorescence analysis of Lif

Fluorescence spectroscopy as a method of sensing distance between two fluorophores provides dynamic insight into the conformational landscape of proteins with high special and temporal resolution. [24] In this work, we used a combination of fluorescence spectroscopy techniques including single-molecule multiparameter fluorescence detection (MFD) in combination with pulsed-interleaved excitation (PIE), in which molecules freely diffuse through the confocal volume and donor and acceptor fluorophores are investigated by separate laser pulses. Burst-wise analysis of individual molecules allows calculating multiple fluorescence observables such as intensity-weighted average fluorescence lifetime $\langle\tau\rangle_F$, anisotropy or FRET efficiency, allowing an investigation of heterogeneities and dynamics of fluorescence species on the sub-millisecond timescale. [25, 26] This is supplemented by an analysis of ensemble FRET-induced

donor decay histograms providing detailed insight into donor-acceptor distance distributions $p(R_{DA})$ of highly dynamic proteins [27] and fluorescence correlation analysis, which expands the accessible time scale of fluorescence dynamics from hundreds of nanoseconds to millisecond. [28]

Lif in solution is characterized by conformational dynamics on sub-millisecond timescale indicated by the "dynamic shift" from the static FRET line (Figure 4A, gray histogram). Static FRET line is a parametric line that links fluorescence observables of a non-exchanging FRET species. [20] Deviations of the histogram from the line towards longer $\langle\tau_D\rangle_F$ values, as in the case of Lif, indicates fast exchange between FRET species within the observation time (ms). Presence of LipA in micromolar concertation and presumably formation of Lif:LipA complex shifts the observed FRET population of Lif towards smaller values of FRET efficiency and towards the static FRET line (Figure 5A, red histogram). The analogous observation was made in case of all FRET variants of Lif (Figure S7)

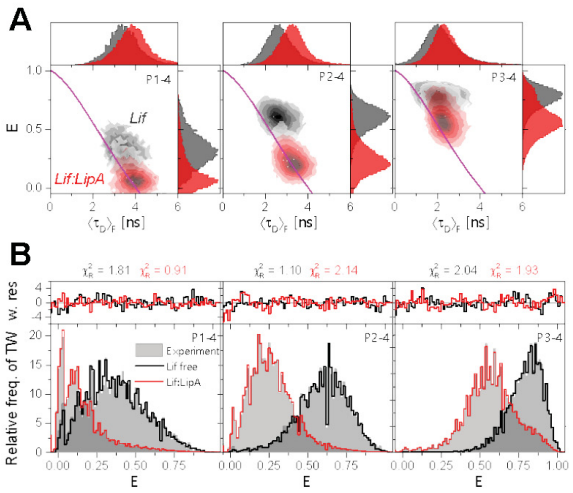


Figure 5. Single-molecule MFD histograms of labeled Lif. (**A**) 2D-MFD analysis of Lif labeled with Alexa488/647 show that free Lif (gray histograms) is on average more compact than Lif in the presence of micromolar concentrations of LipA, indicated by a shift towards lower E values. Magenta line (linker-corrected static FRET line) shows a relation between FRET observables: FRET efficiency E and burst-integrated intensity-

weighted average donor fluorescence lifetime $\langle \tau_D \rangle_F$ for static proteins with a single conformation. The left-side shift of the histogram denotes dynamic exchange of FRET species on the timescale faster than burst duration (milliseconds), prominent in case of free Lif, which is contrasted by Lif:LipA. (**B**) PDA of single-molecule FRET histograms for selected Lif variants shows fits of 1-millisecond FRET averaged donor-acceptor distance $\langle R_{DA} \rangle_E$. The large shift of FRET-populations towards lower FRET (close to static FRET line) is visible for all variants in the presence of a micromolar concentration of LipA, presumably because of Lif:LipA complex formation (K_D approx. 20 nM).

PDA analysis accounts for the broadening of the histograms due to shot noise and distinct acceptor brightness values. Here we used it to extract average FRET distances $\langle R_{DA} \rangle_E$ for Lif and Lif:LipA complexes (Figure 4B). Average FRET distance $\langle R_{DA} \rangle_E$ predicted for homology model [20] disagrees with the measured value of $\langle R_{DA} \rangle_E$ of free Lif in solution in case of all tested FRET variants (Table 2). FRET distances measured in Lif:LipA complex generally agree more with homology model with the exception of variants P1-3 and P2-3. Particularly noteworthy is Lif variant P3-4 with the fluorophores conjugated to the helix α9 present in homology model. Measured FRET distance is shorter than predicted for intact helix α9, which in turn suggests unfolding or substantial bending of the helix, corroborated by MD simulation that showed helix-coil transitions for α9 (Figure 5B).

Table 2 Fluorescence averaged dye distances $\langle R_{DA} \rangle_E$ [29].

Lif variant	Residue 1-2		$\langle R_{DA} \rangle_E$ [29]		
			Homology model[a]	Lif:LipA complex[b]	Lif[b]
P1-2	Q137	A215	59	63 ± 5	58 ± 5
P1-3	Q137	E268	61	50 ± 7	39 ± 5
P1-4	Q137	R296	89	97 ± 7	60 ± 4
P2-3	A215	E268	51	66 ± 6	53 ± 6
P2-4	A215	R296	66	66 ± 6	48 ± 5
P3-4	E268	R296	53	52 ± 6	42 ± 7

[a] FRET averaged donor-acceptor distance simulated with FPS software [20] using homology model of Lif in complex with LipA.
[b] FRET averaged donor-acceptor distance obtained with PDA analysis [30, 31] of smMFD measurements, see figure S8.

We showed that free Lif in solution undergoes a dynamic exchange of states that occur on a sub-millisecond time scale and is on average more compact than Lif:LipA complex. We used ensemble time-correlated single photon-counting (eTCSPC) to obtain FRET-induced donor decay histogram (Figure 6A) and model donor-acceptor distance distributions $p(R_{DA})$ that can resolve conformational heterogeneity of highly dynamic molecules. [27] These distributions

are a combination of fluorophores diffusion restricted by their linkers (typical width of such distribution is approx. 6 – 8 Å) [20] and underlying conformational flexibility of the protein. In the absence of any prior information regarding the number of possible states, we modeled $p(R_{DA})$ with continues generalized Gaussian distribution model with three moments of the distribution as fitting parameters (mean, width, skewness). Example analysis of Lif variant P2-4 shows broad $p(R_{DA})$ with mean inter-dye distance distributions $\langle R_{DA} \rangle$ = 44 Å and width σ_{DA} =18 Å (Figure 6B, gray histogram). In the presence of micromolar LipA, $p(R_{DA})$ shifts toward larger distance and is significantly narrower with $\langle R_{DA} \rangle$ = 63 Å and σ_{DA} = 8 Å (Figure 6B, red histogram) suggesting more open but restricted Lif conformation in complex with LipA. This is in line with MFD and PDA analysis, where Lif exhibits sub-millisecond conformational dynamics and presence of LipA and formation of Lif:LipA complex leads to restriction of dynamics and more extended conformations. Analogous results were observed for all studied variants.

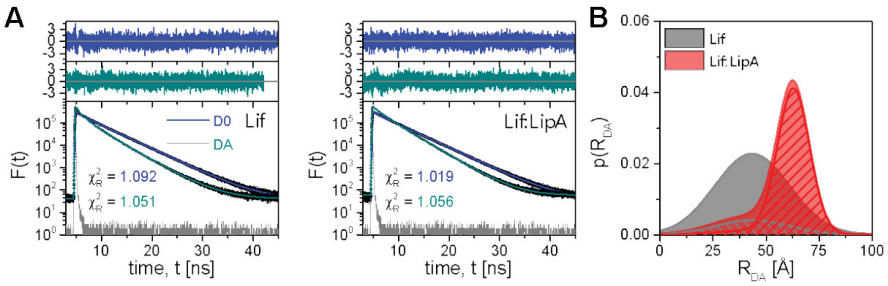


Figure 6. Example of an analysis of the time-resolved fluorescence intensity decay histogram from ensemble TCSPC experiment of Lif variant P2-4. (**A**) Donor-only reference (blue) and donor in presence of acceptor (dark cyan) decay histograms were modeled with 3-exponential decay for donor and generalized Gaussian distribution for the donor in presence of acceptor (Lif) or Gaussian distribution for complex and fraction of free Lif (Lif:LipA). (**B**) Dye-distance distribution $p(R_{DA})$ shows broad distribution of free Lif (gray, $\langle R_{DA} \rangle$ = 44 Å, σ_{DA} =18 Å) suggesting broad range of conformations. Presence of LipA limits the conformations of Lif to a narrow range (Lif:LipA, red, $\langle R_{DA} \rangle$ = 63 Å, σ_{DA} = 8 Å) suggesting restricted Lif conformation in complex with LipA.

Fluorescence correlation analysis of Lif dynamics studied by FRET

Fluorescence correlation techniques use temporal fluctuations in fluorophore brightness, providing insight into the dynamics of FRET exchange, extending the accessible, dynamic

timescales from sub-microseconds to milliseconds. Filter FCS allows to measure the kinetics of exchange between FRET species, when the exchange between two states should result in single anti-correlation time proportional to an inverse of the sum of kinetic rate constants. [28, 32] High- and low-FRET parts of MFD FRET diagrams were used to create filters for FRET species (Figure 7A), and species cross-correlation was computed and fitted with a model containing 3D-Gaussian diffusion and anti-correlation (figure 7B). The anti-correlation relates to the dynamics between FRET species and is shown in figure 7C. Free Lif exhibits a complex character, with anti-correlation ranging from tens of nanoseconds, typical for unstructured peptide chain or dye linker diffusion [33, 34], through microseconds, characteristic for trapped states [35] and fast secondary structure elements folding [36], to hundreds of microseconds, signalizing high energy barriers. [37]

All Lif variants show similar multi-timescale FRET dynamics, extending from nanoseconds to hundreds of microseconds, with the most of FRET dynamics captured by correlation undergoing on the microsecond and faster timescales. Variant P3-4 exhibits increased the amplitude of sub-millisecond time scale dynamics. We also tested four alternative Lif P3*-4 variants: $P3^{255}$-4 (E255C-R296C), $P3^{258}$-4 (E258C-R296C), $P3^{259}$-4 (E259C-R296C) and $P3^{260}$-4 (E260C-R296C). FRET dynamics of P3*-4 is in line with other FRET variants (Figure 7C and S10) with donor-acceptor distance distribution similar to P3-4 variant (Figure 8). We attribute slower dynamics of P3-4 variant to the influence of helix-coil transition of helix 9. It was observed that transition from loop to helix occurs on slower timescale. [38] This is not sensed by P3*-4 FRET variants because of faster motion around the loops connecting helices.

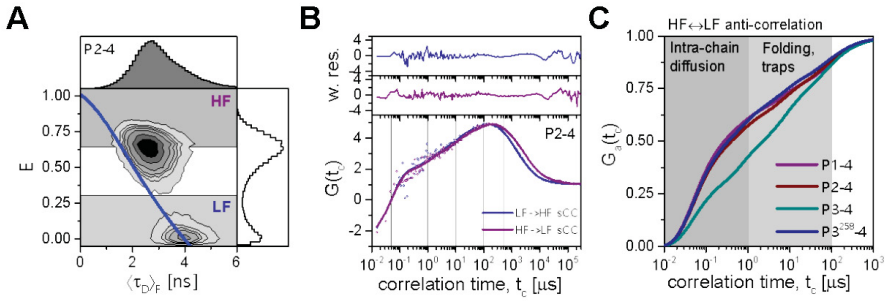


Figure 7. Fluorescence correlation analysis of FRET dynamics. (**A**) Single-molecule MFD histogram of Lif P24 variant. Color boxes denote the selection of bursts for color-lifetime-polarization filters (HF- and LF-species). (**B**) Example of species cross-correlations (LF→HF and HF→LF) created with HF- and LF-filters. Vertical lines show fitted anti-correlation relaxation times $t_{a,i}$. (**C**) Symmetric LF↔HF FRET anti-correlation extracted from species cross-correlation curves (equal terms amplitudes a_i in LF→HF and HF→LF direction). Time-domains assigned to molecular processes, according to Neuweiler *et al.* [39].

Structural ensemble of free Lif

To gain insights at the atomistic level as to the conformational variability of free Lif, we performed an additional 30 independent MD simulations of 250 ns length and pooled the conformations with those of the previous simulations. In order to evaluate the agreement between the computed conformational ensemble and that measured by FRET, we computed the distance distributions for the former using the FPS toolkit (see Materials & Methods section) (figure 8). For the five distances P2-4, $P3^{255}$-4, $P3^{258}$-4, $P3^{259}$-4 and $P3^{260}$-4, the agreement with FRET data is good (χ^2 = 57.9, 49.9, 16.7, 15.3, and 6.4, respectively), and for the three distances P1-2, P1-4, P3-4, a fair agreement is found (χ^2 = 598.3, 128.6, and 463.0, respectively). The largest difference is seen for distances P1-3 and P2-3 (χ^2 = 2619.5 and 2670.7, respectively), which could be due to the high flexibility of the P3 region, as discussed in section 2.4. Still, the overall agreement of the conformational ensemble generated by unbiased MD simulations with experimental data is already remarkable.

To augment the data from the MD simulations with the information captured by FRET on long timescales, we refined the simulated ensemble as described in ref. [40]. Briefly, the MD simulations-generated conformations are clustered with respect to C_α atom root mean square

deviations (C_α-RMSD) with a threshold value of 4 Å. Cluster weights w_k, computed as the size of the cluster normalized to the whole ensemble, are assigned to the respective representative cluster structures. The representative structures of the 545 most populated clusters are pooled as an initial structural ensemble for the refinement, which corresponds to 82% of all structures. The cluster weights are then varied randomly to decrease the discrepancy between simulated and experimental data ($\chi^2_{total} \leq 1$, eq. 10). To control data overfitting, the maximum entropy method (MEM) is used [40, 41] (see Materials & Methods, eq. 11). Hence, to find the optimal set of cluster weights, the free energy function G (eq. 12) is minimized numerically by using simulated annealing employing the temperature-like control parameter θ. For large θ, the change in weights is small and *vice versa*; the best possible agreement with experimental data can be obtained at $\theta = 0$, which allows large changes in the weights. [40] To prevent overfitting, the optimized cluster weights are finally obtained at a relative entropy threshold corresponding to a free energy change of 3 kT (eq. 11, 12) (figure 9A), which is within the expected inaccuracy of MD force fields. [21] The distance distributions computed from the representative ensemble agree with the experimental data to within the estimated error ($\chi^2_{total} \approx 1.8$) (Figure **9**A). As shown in figure 9B, the optimal weights of many clusters obtained at the relative entropy threshold change by multiple orders of magnitude from their initial weights. Most (91 %) weights are down-weighted, while only a 9 % are up-weighted by at most a factor of 100. The representative ensemble contains about 35 distinct conformations of free Lif that jointly contribute to the computed FRET distance distributions. These 35 conformations together constitute about 90 % of the reweighted ensemble (figure 9C). The largest cluster of the representative ensemble has a relative population of 15%. The 35 structures jointly cover the whole range of R_G values (figure S11A).

To probe for robustness, we repeated the refinement two times using different seeds in each case to generate random numbers during the simulated annealing; the averaged optimized

weights ± SEM shown in figure S11B are calculated over the three refinements at the relative entropy threshold. Finally, we also refined the ensemble using a convex optimization approach (COPER). [42] In contrast to MEM, where we limited the ensemble size and weight change, COPER can efficiently reweight all individual populations in a very large ensemble. However, as COPER is not used as a primary method of reweighting in this work, we limit our ensemble to the same 545 clusters. The COPER procedure was then applied to reweight the individual conformations of this ensemble. In COPER method, an optimal solution can be reached more efficiently from an interior point with respect to weights of individual conformations and to find a feasible interior point for the maximum entropy search, an initial χ^2_{total}minimization was carried out. χ^2_{total} was initially minimized to 0.98, starting with the initial weights = 1/N, where N = 545 (see methods for details). The entropy S at the initial point $S(w_k^{equal})$ = ln (545) = 6.3, a significant reduction in entropy value related to a free energy change of -6.8 kT for the minimal χ^2_{total}. Thus, initial minimization of χ^2_{total} led to a very good agreement with the experimental data but also the overfitting of the data. Therefore, the new weights at the minimal χ^2_{total} are taken as the starting point for the maximum entropy search with the same constraints. The subsequent maximum entropy search then yielded data that agree less well with the experimental data than the minimal χ^2_{total} prediction, yet the data is still within the error limits. In turn, the entropy is increased to the threshold corresponding to a free energy change of 3 kT with χ^2_{total} = 1.72. As shown in figure S11C, the optimized weights obtained after MEM and COPER correlate significantly (R^2 = 0.66; $p < 10^{-15}$ according to a significance test for linear regression).

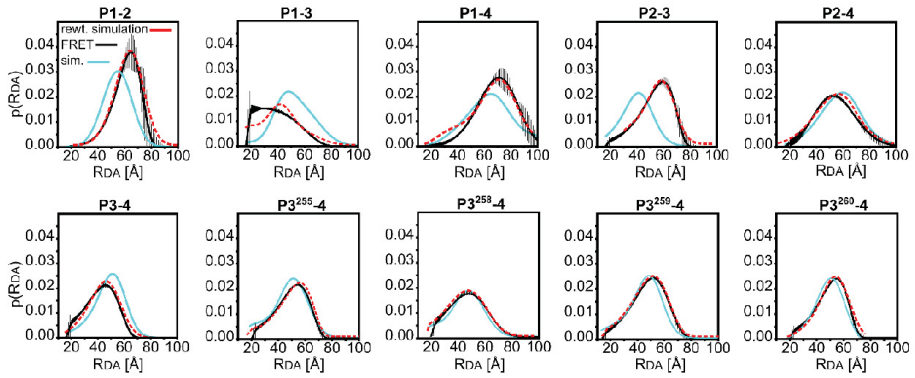


Figure 8. Distance distribution profile obtained for free Lif. The probability of the average donor-acceptor distance ($p(R_{DA})$) is depicted for all ten chosen residue pairs. Distance distributions obtained from eTCSPC analysis are shown in black with the error values as bars, distance distributions obtained from unbiased MD simulations (545 clusters) are shown in cyan color, and from the best refinement run in red. For comparison of $p(R_{DA})$ model fits to time-resolved fluorescence intensity decay histograms see figure S9.

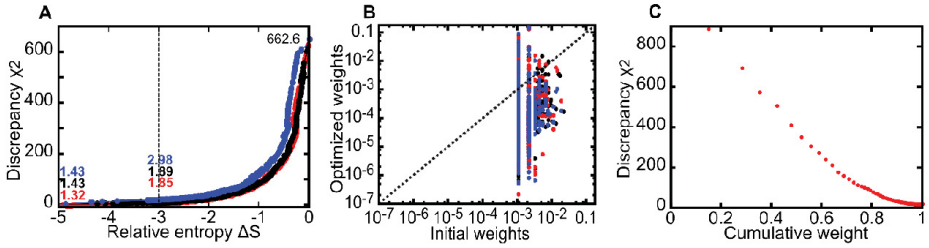


Figure 9. Maximum entropy refinement of cluster weights for free Lif. **(A)** Maximum entropy refinement of cluster weights. The discrepancy χ^2 between experimental and simulated data (eq. 10) *versus* the relative entropy ΔS (eq. 11) is depicted. The black dotted line indicates the relative entropy threshold corresponding to a free energy change of 3 kT below which the simulated data is considered over-fitted. The blue, black, and red curves show three independent runs of refinement. **(B)** Change in initial cluster weights after optimization. The optimized weights are shown at the entropy threshold; the color code is as in panel A. **(C)** 35 most populated clusters after refinement. The discrepancy χ^2 as a function of the cumulative weight of the clusters is plotted for the best refinement run (red curve/dots in panels A/B). Clusters were rank-ordered by their optimized weights.

We further clustered the representative ensemble using the R_G as a cluster criterion with a threshold value of 4 Å, in order to examine the conformations of Lif in the free state. The clustering showed that the representative ensemble is dominated by two clusters. Cluster I (27 structures) contains completely compact conformations of Lif that compact to an R_G of ~21 Å, whereas cluster II (8 structures) contains partially compact conformations of Lif with an R_G of ~25 Å (figure S11A). We further analyzed the similarity of the structures within these two

clusters, by sub-clustering them with respect to C_α-RMSD (threshold value of 3 Å). As shown in figure 10, the structures were similar within the sub-clusters: For cluster I, eight sub-clusters were found, the farthest one having a C_α-RMSD of 5.4 Å from the centroid structure of the initial cluster; for cluster II, four sub-clusters were found, the farthest one having a C_α-RMSD of 5.5 Å from the centroid structure of the initial cluster.

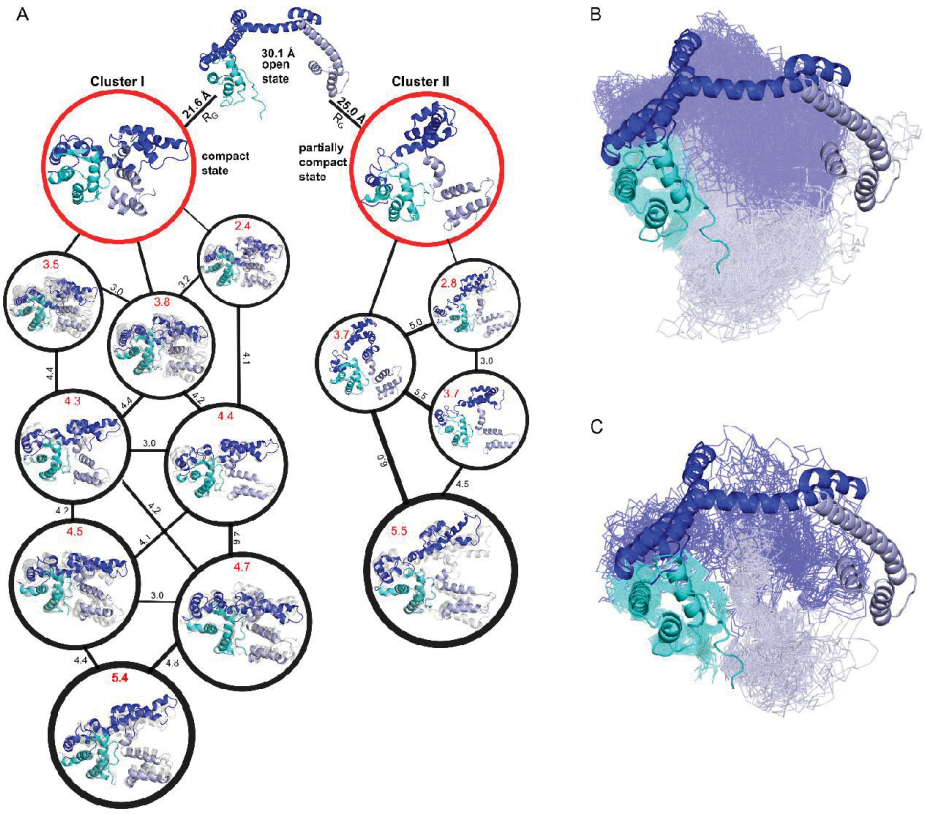


Figure 10. Representative conformational ensemble of free Lif. The top 35 most populated clusters obtained after ensemble refinement by the MEM were first clustered with respect to the R_G with a cut-off of 4 Å; the two resulting clusters I and II (representative structures circled in red) were again clustered with respect to the C_α-RMSD with a cut-off of 3 Å. On the top, the structure of Lif as found in the complex with LipA is shown, with the R_G indicated below. The left branch depicts Lif conformations of cluster I in the "compact state", with the thickness of the circles representing RMSD values of the sub-cluster relative to the representative structure; the RMSD values are also displayed in red within the circles. The right branch depicts Lif conformations of cluster II in the "partially compact state". In addition to the representative conformation color-coded as in figure 2A, further members of the sub-clusters are shown in gray. The thickness of the lines connecting the circles shows the C_α-RMSD between the sub-clusters; this value is also displayed beside the line. **(B)** Overlay of free Lif in the bound conformation (cartoon representation), on the ensemble of 545 cluster representative structures (transparent

ribbon). **(C)** Overlay of free Lif in the bound conformation (cartoon representation), on the ensemble of 35 representative structures found after refinement (transparent ribbon).

Intramolecular interactions in free Lif compared to intermolecular interactions in the Lif:LipA complex

A visual inspection of the representative conformational ensemble of free Lif (figure 10), which contains partially compact and compact conformations, reveals that interactions between the MD1 and MD2 occur. To further understand the nature of the compactness of free Lif, and how such compact Lif still retains the capability to interact with LipA, we computed the probability density distribution of the total number of hydrogen bonds (HBs) and salt bridges (SBs) formed by Lif in the free state as well as in complex with LipA. For the latter, HBs and SBs were determined from conformations of ten independent MD simulations of 1 μs length each. For the former, the reweighted and representatives conformational ensembles of free Lif were used, as was free Lif in the bound conformation.

The total number of HBs and SBs formed by free Lif in the bound conformation decreased upon compaction of the free Lif, which agrees well with the total number of HBs and SBs formed by Lif in the complex (figure 11A, B). Hence, these findings indicate that free, compact Lif minimizes the number of unsatisfied polar interaction sites, resulting in similar numbers of interactions as in the complex. Changes in the solvent-accessible surface area (SASA) of all interface residues as well as hydrophobic residues are at variance with this (figure 11C, D): While the distributions of free, compact LipA show maxima at lower values than that of free Lif in the bound conformation, the distributions only partially overlap with those computed for Lif in complex with LipA. This result demonstrates that free, compact Lif reduces the amount of overall interface SASA as well as hydrophobic SASA; yet, the remaining difference with respect to bound Lif indicates that the Lif:LipA complex should be able to gain affinity from a further burial of, particularly hydrophobic, SASA.

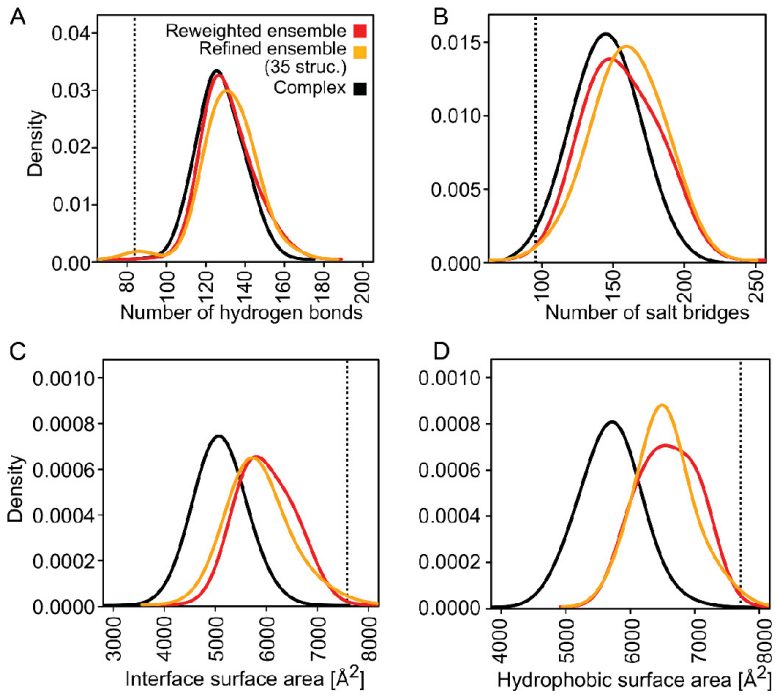


Figure 11. Probability density distributions of the number of total hydrogen bonds (HBs), salt bridges (SBs), solvent-accessible surface area (SASA) of all interface residues and hydrophobic residues for free Lif and Lif in complex with LipA. (A) Probability density of the total number of HBs for the ensemble reweighted by MEM of free Lif (red), the representative conformational ensemble of 35 structures (yellow), and Lif in complex with LipA (black). The black dotted line shows the value of the free Lif in the bound conformation. The distributions of the number of HBs formed by conformations in the reweighted and representatives ensembles agree well with the distribution of HBs formed by Lif in complex with LipA. **(B)** Probability density of the total number of SBs; the color code is as in panel A. Again, all three distributions agree well with each other. **(C)** Probability density distribution of the SASA of interface residues; the color code is as in panel A. The distributions of SASA for all interface residues of the respective ensembles partial overlap with the distribution of bound Lif. **(D)** Probability density distribution of the SASA of hydrophobic residues; the color code is as in panel A. The distributions of SASA of hydrophobic residues of the respective ensembles partial overlap with the distribution of bound Lif.

Discussion

In this study, we have shown by MD simulations and FRET experiments at the atomistic level that the steric chaperone Lif in the absence of its cognate lipase (LipA) does not stay in the open "headphone-like" conformation found for it when bound to LipA, but instead undergoes reversible structural compactions and extensions on a timescale of a millisecond. Furthermore, despite a complete compaction of the "headphone-like" structure in the *apo* state, Lif still

retains the capability to bind to its cognate lipase in that only then hydrophobic and polar interface residues engage completely in interactions, leading to an affinity gain.

Our study was motivated by considerable evidence that showed that Lif is required for the activation of LipA. [5] Over the last years, substantial progress has been made in understanding the mechanism of action of this steric chaperone. [4, 5] Previous *in vitro* experiments indicated that both Lif and LipA from *B. glumae* (*Bg*Lif) undergo conformational rearrangement upon binding, which in turn activates LipA. Furthermore, it was reported that these structural changes are accompanied by the formation of α-helical structures presumably in the extended helical domain (EHD) of Lif. [2] It was also suggested that as all aromatic residues are located within Lif´s mini-domains (MD1 and MD2), therefore, these MDs have appreciable tertiary and secondary structure. [19] Accordingly, here, we analyzed the separated domains of Lif individually with respect to their structural stability. Moreover, we also analyzed the structural dynamics of free Lif by combining molecular simulations with FRET experiments.

First, we performed three independent, unbiased MD simulations of microsecond length in explicit solvent of MD1, MD2, and EHD. The simulations revealed that MD1 retains almost complete the secondary and tertiary structure in the absence of the lipase and other Lif domains. By contrast, EHD (mainly helices 6, 7, and 8) and MD2 (helix 9 and 11) both undergo helix-coil transitions and become less structured when isolated. Second, we performed ten independent unbiased MD simulations of the catalytic folding domain (FOD) of Lif with and without LipA of 1 μs length each. The comparative analysis of these simulations showed that MD2, and to some extent EHD, undergo conformational changes upon binding to EHD: MD2 gains secondary structure, whereas EHD showed both an increase and decrease of helicality in different parts of this domain. By contrast, MD1 does not show any significant change in the secondary structure. To our knowledge, the structural stability of the individual, separated domains of free Lif has not been investigated by molecular simulations before. For the MD

simulations, we used established parameterizations for proteins and solvent, which we had used successfully in other simulations of soluble proteins [43-45]. Furthermore, to consider the impact of a potential force field bias on our results, MD simulations for the isolated Lif domains were performed with the ff99SB [21] and ff14SB [22] force fields. The maximum difference in the results obtained from the two different force field was equivalent to ~1 kT change in free energy, which is within the expected uncertainty of the force fields. [21] As a crystal structure of Lif from *P. aeruginosa* is neither available in the *apo* nor the bound form, a homology model was used as a starting structure for Lif. Note that, as Lif adopts an open "headphone-like" structure in the X-ray structure of lipase:foldase from *B. glumae*, our homology model also has an open, "headphone-like" structure.

Next, the MD simulations of the FOD revealed that in the free form Lif tends to compact, as shown by an R_G of ~21 Å after ~ 200 ns starting from an R_G of 31 Å. However, despite the compactness of the structure, Lif shows considerable internal structural dynamics as monitored by the complete range of the torsion angle occupied (for a definition, see figure 3), revealing a pronounced twisting motion along the EHD that results in multiple conformational states.

To further characterize the structural heterogeneity and dynamics of free Lif in solution, FRET experiments were performed. Six labeled Lif variants were generated and initially tested for *in vitro* efficiency to activate the lipase. Furthermore, the labeled Lif variants were also tested with respect to forming a complex with LipA based on the difference between average FRET efficiency of free Lif and Lif in the complex. The obtained K_D values are more uniform than the measured K_{act}, indicating that all six Lif variants show the same tendency to form complexes with LipA. Next, 2D-MFD analysis of labeled Lif variants shows that free Lif is on average more compact, as indicated by the relatively higher FRET efficiency compared to when LipA is present in μM concentration. This finding confirms the results from unbiased MD simulations, demonstrating the internal consistency of our findings. The left-sided shift of the

MFD histograms indicates the dynamic exchange of FRET species on a timescale of sub-milliseconds, which is prominent in the case of free Lif.

To get further insights into the conformational heterogeneity of free Lif, ensemble time-correlated single photon-counting (eTCSPC) was performed to obtain a FRET-induced donor decay histogram and model donor-acceptor distance distributions $p(R_{DA})$. This analysis revealed that the mean of the inter-dye distance distributions for Lif variants in free Lif shifts to larger values upon addition of LipA. This finding confirmed that in the free state Lif adopts a more compact state as compared to the open state in the complex. The broader distributions for free Lif indicate a higher structural dynamics in that state in contrast to Lif in the complex. This result is again in agreement with the ones above from unbiased MD simulations and MFD analysis.

Next, fluorescence correlation techniques were used to get deeper insights into the dynamics of FRET exchange with respect to accessible dynamic timescales. All six Lif variants showed similar multi-timescale FRET dynamics, extending from nanoseconds to hundreds of microseconds, with most of the FRET dynamics captured by correlations undergoing on the microsecond and faster timescales. Unexpectedly, and so far not fully understood, one Lif variant P3-4 exhibits a decreased amplitude of sub-microsecond time scale dynamics (50 ns). In order to further investigate this finding, four alternative variants (P3*-4) were tested: $P3^{255}$-4 (E255C-R296C), $P3^{258}$-4 (E258C-R296C), $P3^{259}$-4 (E259C-R296C), and $P3^{260}$-4 (E260C-R296C). Interestingly, the four P3*-4 variants showed a FRET dynamics similar to the other Lif variants with donor-acceptor distance distributions similar to the P3-4 variant. Together, the dynamic exchange of FRET species indicates that free Lif reversibly adopts multiple compact and extended conformations that interchange on a timescale from nanoseconds to milliseconds.

Finally, to gain insights at the atomistic level, we reweighted the conformational ensemble of free Lif generated by MD simulations applying experimental constraints from FRET and the

maximum entropy method. Notably, our simulation results agree well with FRET distance distributions even before reweighting, in particular, for P2-4, P3^{255}-4, P3^{258}-4, P3^{259}-4, and P3^{260}-4, P1-2, P1-4, P3-4, with the largest difference being found in the cases of P1-3 and P2-3. After reweighting, our simulation results agree very well with experimental data in terms of FRET distance distributions for all Lif variants. From the reweighted conformational ensemble, 35 representative structures of free Lif were obtained that together represent about 90 % of the reweighted ensemble. These representative structures are dominated by two states of free Lif, a partially compact one with $R_G \sim 25$ Å and a completely compact one with $R_G \sim 21$ Å.

The obtained results open the questions I) why free Lif adopts a compact state and II) how it still retains the capability to bind LipA. As to the former question, the analyses of (changes of) the number of interdomain or intermolecular polar interactions Lif is involved in and of buried solvent-accessible surface area (SASA) of hydrophobic and interface residues revealed that the compaction of free Lif is mainly driven by satisfying polar interactions between Lif domains to the extent also found for the complex with LipA. By contrast, not all hydrophobic and interface residues are buried to the extent found for the complex. Hence, these results suggest Lif tries to minimize disfavorable interactions in the free state. In turn, as to the latter question, these results suggest that the formation of the Lif:LipA complex is driven by a gain in affinity originating from a further burial of remaining hydrophobic and interface residues. In summary, our results shed light onto the structural and functional dynamics of minimalistically structured steric chaperone Lif. Our findings confirmed that in the absence of lipase, Lif tends to adopt multiple conformations ranging from compact to extended conformations, on a timescale vary from nano- to millisecond. In overall, Lif shows to favor a more compact state with a pronounced twisting motion. Although despite compaction of the structure, LipA gains affinity to bind compact Lif via partially exposed hydrophobic and interface surface.

Materials and Methods

Mutation, expression, and purification of Lif and LipA

Lif variants with one Cys mutation were created by standard "QuikChange" (Qiagen) or SLIC [46] PCR methods using Phusion DNA polymerase (Thermo Fischer Scientific), complementary mutagenic oligonucleotide pair (Table S2) and pEHTHis19 plasmid as a template. The resulting mutated plasmids were used as templates for the mutation of a second residue into Cys by the same methods. The presence of desired nucleotide substitutions was confirmed by DNA sequencing.

LipA and Lif wild-type and variants were expressed under the control of the T7 promoter in *E. coli* BL21(DE3) using pETLipA-SS and pEHTHis19 expression plasmids with corresponding mutations, respectively, as described previously. [47]

Lif wild-type and variants carrying an N-terminal His_{10}-tag were purified under native conditions by immobilized metal affinity chromatography using Ni-NTA resins (Qiagen) according to a modified protocol of Hausmann *et al.* [47]. Briefly, *E. coli* BL21(DE3) cells expressing Lif were suspended in buffer containing 50 mM imidazole, disrupted by a French Press, and centrifuged at 40,000 g for 30 min. Cell lysate containing Lif was subjected to the Ni-NTA column followed by washing and elution with buffers containing 50 mM and 500 mM imidazole, respectively. Lif was transferred into Tris-HCl buffer (50 mM, pH 8, 1 mM EDTA) using an ultrafiltration device (Vivaspin) with a membrane of 10 kDa pore size and stored at -20°C.

Cells expressing insoluble inclusion bodies of LipA were resuspended in Tris-HCl buffer (100 mM, pH 8) containing 5 mM EDTA and disrupted by a French Press. LipA inclusion bodies were collected by centrifugation at 10,000 g for ten min and suspended in the same buffer. Centrifugation and wash steps were repeated three times to obtain purified LipA inclusion bodies, which were solubilized with Tris-HCl buffer (100 mM, pH 8) containing 8 M urea at

37°C for 1 h. Remaining insoluble material was removed by centrifugation at 10,000 *g* for ten min.

Biochemical analysis of wild-type Lif and FRET Lif variants

***In vitro* activation of LipA with Lif.** For activation, LipA inclusion bodies were solubilized with Tris-HCl buffer (100 mM, pH 8) containing 8 M urea at 37°C for 1 h. The remaining insoluble material was removed by centrifugation at 10,000 g for ten min. Solubilized LipA (68 µM) was diluted tenfold with refolding buffer (10 mM Tris-HCl, 1 mM $CaCl_2$, 45% (v/v) glycerol, pH 8) containing an equimolar amount of Lif and was gently agitated for 3 h at 30°C.

LipA activity assay. The activity of LipA was spectrophotometrically monitored by the release of *p*-nitrophenolate from the standard lipase substrate *p*-nitrophenyl palmitate. [15]

Fluorescence labeling of Lif double Cys mutants

For fluorescence studies, cysteine residues of Lif variants were modified with malemide functionalized dyes, Alexa Fluor 488 C5 maleimide (donor) and Alexa Fluor 647 C2 maleimide (acceptor) (ThermoFisher Scientific), in two subsequent steps. For that purpose 100 nmol of purified Lif in labeling buffer (50 mM Tris-HCl pH 8.0 at 4°C), 0.8 molar parts of Alexa488 (ε_{488nm} = 71,000 M^{-1} cm^{-1}) dissolved in dimethyl sulfoxide (DMSO) and 5 mM tris(2-carboxyethyl) phosphine (pH 8) were agitated slowly for 3 h at 25°C. Theoretical extinction coefficient of double Cys mutants of Lif of ε_{280nm} = 19,940 M^{-1} cm^{-1} was used to calculate the Lif concentrations of double-labeled Lif variants according the equation: c(Lif) / (mg/ml) = A_{280nm} - (0.11 x A_{488nm}) - (0.02 x A_{647nm}) / ε_{280nm}(Lif). Single donor labeled Lif was purified at 10°C *via* anion exchange chromatography using ResourceQ column (1 ml capacity) (GE Healthcare) according to the manufacturer recommendations. Briefly, donor labeled Lif sample diluted five times in IEX-A buffer (50 mM Bis-Tris pH 6.3 at 10°C) were loaded to the ResourceQ column followed by a wash step with 10 mL of IEX-A Buffer. Elution of proteins bound to the column was achieved with a linear increase of NaCl up to 500 mM during 180

min. Fractions of single labeled Lif with protein to donor dye ratio of 1:1 were concentrated, and the buffer was exchanged to the labeling buffer using an ultrafiltration device (Vivaspin) with a membrane of 10 kDa pore size. Labeling with Alexa647 dye and purification of double-labeled Lif variants was performed following the donor labeling procedure using four-fold molar excess of the acceptor dye. Labeled Lif samples were stored at -20°C in the presence of 20% (v/v) glycerol. For control fluorescence experiments single Cys mutants of Lif were labeled with Alexa488 and Alexa647 dye separately as described for double Cys mutants only that second labeling step was omitted.

Single-molecule Multiparameter Fluorescence Detection (MFD).

Data collection. MFD measurements [25, 48, 49]was done using a 485 nm diode laser (LDH-D-C 485 PicoQuant, Germany, operating at 32 MHz, power at objective 40 μW) and 635 nm diode laser (PicoQuant, Germany, operating at 32 MHz, power at objective 8 μW) in interleaved pulse mode, PIE [26] exciting freely diffusing labeled molecules that passed through a detection volume of the 60X, 1.2 NA collar (0.17) corrected Olympus objective. To ensure the single-molecule conditions, confocal volume occupancy was kept below 5%. The emitted fluorescence signal was collected through the same objective and spatially filtered using a 100 μm pinhole, to define an effective confocal detection volume. Then, the signal was divided into parallel and perpendicular components at two different colors (“green” and “red”) through band pass filters, HQ 520/35 and HQ 720/150, for green and red respectively, and split further with 50/50 beam splitters. In total, eight photon-detectors are used- four for green (τ-SPAD, PicoQuant, Germany) and four for red channels (APD SPCM-AQR-14, Perkin Elmer, Germany). HydraHarp 400 Multichannel Picosecond Event Timer & TCSPC Module (PicoQuant, Germany) was used for data registration. Measurements were performed at room temperature (to 21°C±1°C).

Data analysis. Single-molecule Multiparameter Fluorescence Detection data were processed as described in ref. [25] Fluorescence intensity-based photon distribution analysis, PDA was performed according to refs. [30, 31] using the model of single Gaussian state with variable width for free Lif and a model of with two Gaussian was used for Lif in the presence of micromolar LipA. Uncertainty due to orientation factor κ^2 was treated as described in Sindbert *et al.* [50]. Correlation analysis, filtered FCS was preformed, according to Felekyan *et al.* [28]. Because a large number of terms was required to represent the data in global fit of species auto- and cross-correlations, a formal model with four anti-bunching terms with fixed relaxation times (50 ns, 1 µs, 10 µs, 100 µs) was used to idealize the anti-correlation part of species cross-correlation curves for the comparison.

$G(t_c) = 1 + \frac{1}{N} G_{dif}(t_c) \times G_b(t_c) \times G_{ph}(t_c) \times G_a(t_c)$, where

$$G_{dif}(t_c) = \left(1 + \frac{t_c}{t_{dif}}\right)^{-1} \left(1 + \left(\frac{\omega_0}{z_0}\right)^2 \times \frac{t_c}{t_{dif}}\right)^{-1/2}$$

$$G_b(t_c) = 1 - b + b\ exp\left(\frac{-t_c}{t_b}\right)$$

$$G_{ph}(t_c) = \left(1 - a_{ph}\ exp\left(\frac{-t_c}{t_{ph}}\right)\right)$$

$$G_a(t_c) = \left(1 - a \times \sum_{i=1}^{4} a_i\ exp\left(\frac{-t_c}{t_{a,i}}\right)\right) \qquad \text{Eq. 1}$$

The observation volume is approximated by a 3D-Gaussian volume with $1/e^2$ radii in the lateral (ω_0) and the axial direction (z_0), t_{dif} is the diffusion time, b and t_b is the amplitude and time of the bunching term, a_{ph} and t_{ph} is the amplitude and time of the anti-bunching term related to acceptor photo-bleaching, and a, a_i and $t_{a,i}$ are the amplitudes and times of the anti-correlation terms.

Determination of Acceptor Quantum Yield. Acceptor fluorescence quantum yield $\Phi_{F,A}$ was determined via *PIE-MFD* [26] for each Lif FRET variant. The average fluorescence lifetime of the acceptor $\langle\tau_{(A|A)}\rangle_x$ was determined via sub-ensemble analysis of fluorescence intensity

decay of the acceptor dye present on FRET molecules. TAC-gated (direct acceptor excitation channel) fluorescence correlation curves were used to determine the fraction of the dark (non-fluorescence states including triplet and cis-trans isomerization) a_{dark}. Quantum yield was determined using reference value for AlexaFluor647: $\Phi_{F,A,\ ref} = 0.4$ for $\langle \tau_{A\,|\,A} \rangle_{x,ref} = 1$ ns [51] using:

$$\Phi_{F,A} = 0.4 \cdot \langle \tau_{(A|A)} \rangle_x \cdot a_{dark} \qquad \text{Eq. 2}$$

Lif:LipA complex titration. Labeled Lif was mixed with a series of concentrations of LipA in 5 mM Tris, 5 mM glycine buffer at pH 9 (final concentration of Lif: 5 nM), and incubated overnight at room temperature. Protein low-binding labware, protein low-binding tubes (Eppendorf) and 200 µl and 10 µl pipette ultra-low retention tips (Brand), was used to minimize loses of LipA due to protein adsorption. Measurements were performed on a single molecule setup described above using 485 nm diode laser (operating at 64 MHz, power at objective 120 µW).“Green” and “red” signal, S_G and S_R, was converted into a proximity ratio $P = S_R / (S_G + S_R)$. P of free Lif was set to the fraction of complex $x_{complex} = 0$ and P of Lif at 1.5 µM LipA concentration to $x_{complex} = 1$. Data were fitted to Eq. 2 using OriginPro9 (OriginLab Corp., USA):

$$x_{complex} = \frac{c_{LipA}}{K_D + c_{LipA}} \qquad \text{Eq. 3}$$

Ensemble time-correlated single photon-counting (eTCSPC).

Data collection. Ensemble time-correlated single photon-counting (eTCSPC) measurements of high-precision fluorescence intensity decays were performed on a FluoTime300 High-Performance Fluorescence Lifetime Spectrometer equipped with HydraHarp 400 Multichannel Picosecond Event Timer & TCSPC Module and Hybrid Photomultiplier Detector Assembly (Picoquant GmbH, Germany). The excitation source was a SuperK EXTREME supercontinuum white light laser (NKT Photonics, Denmark), set to 488 nm and 19.51 MHz

pulsing frequency, passed through BP488/10x (Chroma Technology, USA) interference filter for excitation beam cleaning. The emission wavelength was set to 525 nm, and the long-pass LP495nm filter was used to reduce the scattered light contribution. All measurements were performed using magic-angle polarizer configuration (54.7°) and with temperature control set to 20°C±0.1°C. Time-resolved fluorescence intensity decays were collected using 8ps bins and analyzed using software for analysis of fluorescence data *ChiSurf* [27].

Data analysis. FRET efficiency distribution $p(E)$ or equivalent inter-dye distance distribution $p(R_{DA})$ can be determined by analyzing the FRET-induced fluorescence intensity decay of a donor dye $F_{D(A)}(t)$, when the fluorescence properties of the donor in the absence of FRET $F_{D(0)}(t)$ modeled as a sum of exponentials with species fraction x_i and lifetime τ_i, are known [27].

$$F_{D(0)}(t) = \sum_{i=1}^{3} x_i \cdot e^{-t/\tau_i}$$

$$F_{D(A)}(t) = F_{D(0)}(t) \int_{R_{DA}} p(R_{DA}) \exp\left(-tk_0(R_0/R_{DA})^6\right) dR_{DA} + x_{D(0)} F_{D(0)}(t) \qquad \text{Eq. 4}$$

R_0 is the Förster radius (here R_0 = 52 Å), k_0 is the radiative rate of unquenched donor dye.

To model the donor-acceptor distance distribution we used Gaussian function (model parameters: number of Gaussian states n, state fraction x_i, state mean inter-dye distance mean $\langle R_{DA}\rangle_i$, distance distribution width σ_{DAi}):

$$p(R_{DA}) = \sum_{i=1}^{n} x_i \frac{1}{\sqrt{2\pi\sigma_{DAi}^2}} \exp\left(-\frac{(R_{DA}-\langle R_{DA}\rangle_i)^2}{2\sigma_{DAi}^2}\right) \qquad \text{Eq. 5}$$

or generalized single Gaussian function (model parameters: mean $\langle R_{DA}\rangle$, width σ_{DA}, shape factor κ):

$$p(R_{DA}) = \frac{1}{\sigma_{DA} - \kappa(R_{DA}-\langle R_{DA}\rangle)} \frac{1}{\sqrt{2\pi\sigma_{DA}^2}} \exp\left(-\frac{\left(-\frac{1}{\kappa} log\left[1-\frac{\kappa(R_{DA}-\langle R_{DA}\rangle)}{\sigma_{DA}}\right]-\langle R_{DA}\rangle\right)^2}{2\sigma_{DA}^2}\right) \qquad \text{Eq. 6}$$

Molecular dynamics simulations

All-atom MD simulations were performed with the Amber11 software package [52], using the ff99SB and ff14SB force fields [21] as described by Ciglia et at.[45] Lif in complex with LipA, free Lif and the separated domains of free Lif, MD1, EHD, and MD2, were placed in octahedral periodic boxes of TIP3P water molecules [43] such that the smallest distance between the edges of the box and the closest solute atom is 11 Å. The SHAKE algorithm [53] was applied to constrain bond lengths of hydrogen atoms, and long-range electrostatic interactions were taken into account using the Particle Mesh Ewald method [54]. The time step was set to 2 fs with a non-bonded cut-off of 8 Å. The starting structures were first energy minimized by applying 50 steps of steepest descent minimization, followed by 450 steps of conjugate gradient minimization. During the minimization, the solute atoms were restrained applying decreasing harmonic potentials, with force constant of 25 kcal mol^{-1} $Å^{-2}$ initially, reduced to 5 kcal mol^{-1} $Å^{-2}$ in the last round. For thermalization, the systems were heated from 100 K to 300 K in 50 ps of canonical (NVT)–MD simulations with force constant of 5 kcal mol^{-1} $Å^{-2}$. Afterward, MD simulations of 250 ps length were performed using the isothermal-isobaric ensemble (NPT)-MD simulations with the same force constant. Finally, the force constant of the harmonic restraints was reduced to zero during 100 ps of MD simulations in the NVT ensemble. For production, 10 MD simulations of 1 μs length for free Lif as well as for the Lif:LipA complex were performed and for each domain three MD simulations of 1 μs length each. 30 additional MD simulations of 250 ns length for free Lif were also performed to generate a conformational ensemble.

Analysis of trajectories

MD trajectories were analyzed with the Amber module CPPTRAJ [55] in terms of the average α-helix content, a dihedral angle to describe twist motions (figure 2 and 3) and compactness

(R_G). The ten MD simulations of 1 μs length each were used for these analyses; the first 100 ns of these trajectories were not used in the R_G analysis. The average α-helix content per-residue was expressed as the difference between bound Lif with respect to free Lif and analyzed using the two-sided *t*-test. A *p*-value < 0.05 was considered statistically significant (figure 2C). Similarly, average α-helix contents per-residue were calculated for individual domains over three independent MD simulations. The centers of mass of residues Q137, A215, E268, and R296 were used to define the dihedral angle (figure 3A, B). Next, the total number of hydrogen bonds (HBs) and salt bridges (SBs) formed by free Lif were also calculated by using CPPTRAJ, over the conformations of the reweighted ensemble (545 structures) and the representative ensemble (35 structures). For the Lif in complex with LipA, all conformations generated during ten independent MD simulations of Lif:LipA complex were used to calculate a total number of HBs and SBs. The density distributions were computed by using kernel density estimator. Similarly, density distributions for total hydrophobic solvent-accessible surface area (SASA) and the interface SASA were calculated for free Lif and Lif in complex with LipA.

FRET-accessible volume calculations

To compare data from MD simulations with FRET results, we produced accessible volume (AV) clouds [50] for all MD-generated conformations of free Lif and then calculated the distance between the mean positions of the dyes attached at defined residues (Table 1). In detail, the AVs of dye molecules attached to the chosen residues Q137, A215, E268, E255, R258, Q259, M260, and R296 via linkers in free Lif conformations were computed using the FPSv1.2 program [20]. The algorithm implemented in FPSv1.2 uses a coarse-grained representation in which the label is approximated by a flexibly-linked sphere to compute the sterically allowed AV [24]. Here, the linkers of Alexa488 and Alexa647 are approximated as flexible tubes with a width of L_{width} = 4.5 Å and a length of L_{link} = 20.5 Å and 21.0 Å, respectively. One radii was used to describe the dyes (5.0 Å, 4.5 Å, 1.5 Å) and (11.0 Å, 4.7 Å, 1.5 Å) for Alexa488 and

Alexa647, respectively. From the accessible volumes, FPS computes $\langle R_{DA} \rangle$, the mean donor-acceptor distance, and $\langle R_{DA} \rangle_E$, the FRET averaged mean donor-acceptor distance, for each pair of AVs.

Cluster reweighting

To generate an initial ensemble of conformations for reweighting, the MD-generated conformations of free Lif were clustered based on their structural similarity. Cluster analysis was performed with the Amber module CPPTRAJ [55] using the hierarchical agglomerative algorithm. As a distance measure, the C_α atom root mean square deviation (Cα-RMSD) of all residues was used. A maximal distance between all members of two clusters (complete linkage) of < 4 Å was used as ending criterion for the clustering. With these settings, we obtained a total of 1510 clusters of free Lif structures. About 64 % of all clusters contained only a single structure. The representative structures of the other 545 clusters (i.e., the structure closest to the centroid of each cluster) were pooled together to generate an initial structural ensemble for refinement. This ensemble corresponds to 82% of all structures. Cluster weights w were calculated by computing the size of the cluster as a fraction of all structures.

Cluster reweighting was performed similarly as described in Rozycki (2011) [40]. First, we calculated the Gaussian probability distance distribution with mean value $\langle R_{DA} \rangle$ and standard deviation σ_{DA} (eq. 7) for all dye pairs in a structure [50].

$$p(R_{DA}) = \frac{1}{\sqrt{2\pi\sigma_{DA}{}^2}} \exp\left(-\frac{(R_{DA}-\langle R_{DA}\rangle)^2}{2\sigma_{DA}{}^2}\right) \quad \text{Eq. 7}$$

The average $p(R_{DA})$ for the whole ensemble of simulated structures is computed according to eq. 8, where N is the total number of clusters, w_k is the weight of cluster k, normalized to $\sum_k w_k = 1$, and $p(R_{DA})_k$ is the mean probability distance distribution for cluster k.

$$p(R_{DA})_{sim} = \sum_{k=1}^{N} w_k (p(R_{DA})_k) \quad \text{Eq. 8}$$

The discrepancy between the simulated, average $p(R_{DA})_{sim}$ and experimental $p(R_{DA})_{exp}$ (Table S9) for any dye pair (*i*, *j*) was calculated using eq. 9

$$\chi^2_{(i,j)} = \frac{1}{N_b}\sum_{m=1}^{N_b}\frac{\left(p(R_{DA})_{(i,j)sim}(m) - p(R_{DA})_{(i,j)exp}(m)\right)^2}{\sigma^2_{DA(i,j)}(m)} \quad \text{Eq. 9}$$

N_b is the total number of bins in a distance distribution, and $\sigma^2_{DA}(m)$ is the statistical error calculated from FRET experiments for bin *m*. The total deviation (χ^2_{total}) is calculated by summing up the calculated χ^2 for all dye pairs according to eq. 10

$$\chi^2_{total} = \sum \chi^2 \quad \text{Eq. 10}$$

In order to improve the agreement with experimental data, we vary the cluster weights w_k randomly in each iteration. This procedure was repeated until $\chi^2_{total} \leq 1$. To prevent the overfitting of data, we used the maximum entropy method (MEM) [41], for which we defined the relative entropy (eq. 11) as described in Rozycki (2011) [40].

$$\Delta S = -\sum_{k=1}^{N} w_k \ln\frac{w_k}{w_k^{(0)}} \quad \text{Eq. 11,}$$

where $w_k^{(0)}$ is the initial cluster weight, and w_k is the varied weight. Hence, at the start, when $w_k = w_k^{(0)}$, $\Delta S = 0$.

Next, to obtain an optimal solution, we used the same free energy function *G* (eq. 12) introduced in ref. [40] with a temperature-like control parameter θ.

$$G = \chi^2 - \theta\Delta S \quad \text{Eq. 12}$$

The function *G* was minimized by using simulated annealing with respect to the normalized weights of the clusters. For a large value of θ, minimizing *G* leads to small perturbations of w_k from the initial values $w_k^{(0)}$ for the majority of the clusters [40]. In contrast, minimizing *G* with

θ close to zero produces the best agreement with experiment, but also causes large changes in the weights, which may a result of overfitting. Therefore, the simulated annealing was started at a large value of θ = ~ 700 to allow minimal changes in initial weights, which was then reduced by a factor of 0.9 in repeated fixed steps to sweep out a broad θ range [56]. G was minimized until the statistical uncertainty $\chi^2_{total} \approx 1$, and no further decrease in χ^2_{total} was obtained over the iterations. We chose the ΔS threshold corresponding to a free energy change of 3 kT, below which we considered the data over-fitted, as this value is within the expected error of MD force fields.

After reweighting, we rank-ordered the clusters by their new weights in decreasing order, calculated the discrepancy χ^2 and plotted it as a function of the cumulative new weights (Figure 8C). Top 35 most populated clusters account 90 % for the reweighted ensemble. The ensemble of 35 representative structures was clustered, first, by using the R_G as a criterion with a distance cut-off of 4 Å as we used before, between members of two clusters. The two resulting clusters were sub-clustered using the C_α-RMSD as a criterion with a distance cut-off of 3 Å between the structures of the same cluster in order to analyze the structural similarity within the clusters. Both clusterings were performed using CPPTRAJ, as explained above.

The MEM refinement was repeated two more times, with different seeds for the random number generator during simulated annealing for G minimization. Therefore, in total, we ran the refinement three times starting from the same initial weight distribution. From the three runs, average optimized weights ± SEM for each cluster were computed at the relative entropy threshold of 3kT. However, in order to check the efficiency of MEM to provide the optimal solution, we reweighted the simulated ensemble of 545 clusters using the convex optimization ensemble reweighting (COPER) method as described in Leung (2016) [42]. In brief, this approach uses an interior point method [57] as a starting point to effectively find the optimal solution using the same MEM as explained above. An interior point is a point where the

experimental constraints meet the best solution, that is, $\chi^2_{total} \leq 1$. We found the interior point by minimizing χ^2_{total} with the free energy function $G = kT\,\Delta S$, where ΔS is defined in eq. 11, k is the Boltzmann constant, and T is the absolute temperature. As G is minimized without the control parameter, we obtained the minimal value of $\chi^2_{total}$with a large change in the ΔS, which indicates data overfitting. The weights obtained at the minimal χ^2_{total} are taken as the starting weights to perform the refinement again in the inverse way as compared to MEM to maximize the relative entropy. The refinement was done until the relative entropy threshold was reached.

References

1. Ellis, R.J. and Vandervies, S.M., *Molecular Chaperones.* Annual Review of Biochemistry, 1991. **60**: p. 321-347.
2. El Khattabi, M., Van Gelder, P., *et al.*, *Role of the lipase-specific foldase of Burkholderia glumae as a steric chaperone.* Journal of Biological Chemistry, 2000. **275**(35): p. 26885-26891.
3. Jaeger, K.E., Ransac, S., *et al.*, *Bacterial Lipases.* Fems Microbiology Reviews, 1994. **15**(1): p. 29-63.
4. Rosenau, F., Tommassen, J., *et al.*, *Lipase-specific foldases.* Chembiochem, 2004. **5**(2): p. 153-161.
5. Pauwels, K., del Pino, M.M.S., *et al.*, *Decoding the Folding of Burkholderia glumae Lipase: Folding Intermediates En Route to Kinetic Stability.* Plos One, 2012. **7**(5).
6. Pauwels, K., Van Molle, I., *et al.*, *Chaperoning Anfinsen: the steric foldases.* Molecular microbiology, 2007. **64**(4): p. 917-922.
7. Ihara, F., Okamoto, I., *et al.*, *Lipase Modulator Protein (Liml) of Pseudomonas Sp Strain-109.* Journal of Bacteriology, 1995. **177**(5): p. 1254-1258.
8. Douzi, B., Ball, G., *et al.*, *Deciphering the Xcp Pseudomonas aeruginosa Type II Secretion Machinery through Multiple Interactions with Substrates.* Journal of Biological Chemistry, 2011. **286**(47): p. 40792-40801.
9. Konig, B., Jaeger, K.E., *et al.*, *Role of Pseudomonas aeruginosa lipase in inflammatory mediator release from human inflammatory effector cells (Platelets, granulocytes, and monocytes).* Infection and Immunity, 1996. **64**(8): p. 3252-3258.
10. Kang, Y., Zarzycki-Siek, J., *et al.*, *Multiple FadD Acyl-CoA Synthetases Contribute to Differential Fatty Acid Degradation and Virulence in Pseudomonas aeruginosa.* Plos One, 2010. **5**(10).
11. Henzler-Wildman, K. and Kern, D., *Dynamic personalities of proteins.* Nature, 2007. **450**: p. 964.
12. Funken, H., Knapp, A., *et al.*, *The Lipase LipA (PA2862) but Not LipC (PA4813) from Pseudomonas aeruginosa Influences Regulation of Pyoverdine Production and Expression of the Sigma Factor PvdS.* Journal of Bacteriology, 2011. **193**(20): p. 5858-5860.
13. Jaeger, K., *Extracellular enzymes of Pseudomonas aeruginosa as virulence factors.* Immunitat und Infektion, 1994. **22**(5): p. 177-180.
14. Frenken, L.G.J., Bos, J.W., *et al.*, *An Accessory Gene, Lipb, Required for the Production of Active Pseudomonas-Glumae Lipase.* Molecular Microbiology, 1993. **9**(3): p. 579-589.

15. Jaeger, K.E. and Kovacic, F., *Determination of lipolytic enzyme activities.* Methods in Molecular Biology, 2014. **1149**: p. 111-34.
16. El Khattabi, M., Ockhuijsen, C., *et al.*, *Specificity of the lipase-specific foldases of gram-negative bacteria and the role of the membrane anchor.* Molecular and General Genetics, 1999. **261**(4-5): p. 770-776.
17. Shibata, H., Kato, H., *et al.*, *Molecular properties and activity of amino-terminal truncated forms of lipase activator protein.* Bioscience Biotechnology and Biochemistry, 1998. **62**(2): p. 354-357.
18. Filloux, A., *The underlying mechanisms of type II protein secretion.* Biochimica Et Biophysica Acta-Molecular Cell Research, 2004. **1694**(1-3): p. 163-179.
19. Pauwels, K., Lustig, A., *et al.*, *Structure of a membrane-based steric chaperone in complex with its lipase substrate.* Nature structural & molecular biology, 2006. **13**(4): p. 374-5.
20. Kalinin, S., Peulen, T., *et al.*, *A toolkit and benchmark study for FRET-restrained high-precision structural modeling.* Nature Methods, 2012. **9**(12): p. 1218-U129.
21. Hornak, V., Abel, R., *et al.*, *Comparison of multiple Amber force fields and development of improved protein backbone parameters.* Proteins: Structure, Function, and Bioinformatics, 2006. **65**(3): p. 712-725.
22. Maier, J.A., Martinez, C., *et al.*, *ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB.* Journal of Chemical Theory and Computation, 2015. **11**(8): p. 3696-3713.
23. Adamo, R., Sokol, S., *et al.*, *Pseudomonas aeruginosa Flagella Activate Airway Epithelial Cells through asialoGM1 and Toll-Like Receptor 2 as well as Toll-Like Receptor 5.* American Journal of Respiratory Cell and Molecular Biology, 2004. **30**(5): p. 627-634.
24. Dimura, M., Peulen, T.O., *et al.*, *Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems.* Current Opinion in Structural Biology, 2016. **40**: p. 163-185.
25. Sisamakis, E., Valeri, A., *et al.*, *Accurate Single-Molecule Fret Studies Using Multiparameter Fluorescence Detection.* Methods in Enzymology, Vol 475: Single Molecule Tools, Pt B, 2010. **475**: p. 455-514.
26. Kudryavtsev, V., Sikor, M., *et al.*, *Combining MFD and PIE for Accurate Single-Pair Forster Resonance Energy Transfer Measurements.* Chemphyschem, 2012. **13**(4): p. 1060-1078.
27. Peulen, T.O., Opanasyuk, O., *et al.*, *Combining Graphical and Analytical Methods with Molecular Simulations To Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately.* Journal of Physical Chemistry B, 2017. **121**(35): p. 8211-8241.
28. Felekyan, S., Kalinin, S., *et al.*, *Filtered FCS: Species Auto- and Cross-Correlation Functions Highlight Binding and Dynamics in Biomolecules.* Chemphyschem, 2012. **13**(4): p. 1036-1053.
29. Tang, Y.-W., *et al.*, *Index*, Molecular Medical Microbiology (Second Edition), 2015, Academic Press: Boston. p. 2097-2145.
30. Antonik, M., Felekyan, S., *et al.*, *Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis.* Journal of Physical Chemistry B, 2006. **110**(13): p. 6970-6978.
31. Kalinin, S., Felekyan, S., *et al.*, *Probability distribution analysis of single-molecule fluorescence anisotropy and resonance energy transfer.* Journal of Physical Chemistry B, 2007. **111**(34): p. 10253-10262.
32. Felekyan, S., Sanabria, H., *et al.*, *Analyzing Forster Resonance Energy Transfer with Fluctuation Algorithms.* Fluorescence Fluctuation Spectroscopy (Ffs), Pt B, 2013. **519**: p. 39-85.
33. Moglich, A., Joder, K., *et al.*, *End-to-end distance distributions and intrachain diffusion constants in unfolded polypeptide chains indicate intramolecular hydrogen bond formation.*

Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(33): p. 12394-12399.

34. Nettels, D., Hoffmann, A., *et al.*, *Unfolded Protein and Peptide Dynamics Investigated with Single-Molecule FRET and Correlation Spectroscopy from Picoseconds to Seconds (vol 112, pg 6137, 2008)*. Journal of Physical Chemistry B, 2012. **116**(19): p. 5801-5801.
35. Yirdaw, R.B. and Mchaourab, H.S., *Direct Observation of T4 Lysozyme Hinge-Bending Motion by Fluorescence Correlation Spectroscopy*. Biophysical Journal, 2012. **103**(7): p. 1525-1536.
36. Chung, H.S., McHale, K., *et al.*, *Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times*. Science, 2012. **335**(6071): p. 981-984.
37. Kilic, S., Felekyan, S., *et al.*, *Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1 alpha*. Nature Communications, 2018. **9**.
38. Chung, H.S., Louis, J.M., *et al.*, *Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(29): p. 11837-11844.
39. Neuweiler, H., Johnson, C.M., *et al.*, *Direct observation of ultrafast folding and denatured state dynamics in single protein molecules*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(44): p. 18569-18574.
40. Rozycki, B., Kim, Y.C., *et al.*, *SAXS Ensemble Refinement of ESCRT-III CHMP3 Conformational Transitions*. Structure, 2011. **19**(1): p. 109-116.
41. Jaynes, E.T., *Information Theory and Statistical Mechanics*. Physical Review, 1957. **106**(4): p. 620-630.
42. Leung, H.T.A., Bignucolo, O., *et al.*, *A Rigorous and Efficient Method To Reweight Very Large Conformational Ensembles Using Average Experimental Data and To Determine Their Relative Information Content*. Journal of Chemical Theory and Computation, 2016. **12**(1): p. 383-394.
43. Jorgensen, W.L., Chandrasekhar, J., *et al.*, *Comparison of Simple Potential Functions for Simulating Liquid Water*. Journal of Chemical Physics, 1983. **79**(2): p. 926-935.
44. Pagani, G. and Gohlke, H., *On the contributing role of the transmembrane domain for subunit-specific sensitivity of integrin activation*. Scientific reports, 2018. **8**(1): p. 5733.
45. Ciglia, E., Vergin, J., *et al.*, *Resolving Hot Spots in the C-Terminal Dimerization Domain that Determine the Stability of the Molecular Chaperone Hsp90*. Plos One, 2014. **9**(4).
46. Li, M.Z. and Elledge, S.J., *Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC*. Nature Methods, 2007. **4**(3): p. 251-256.
47. Hausmann, S., Wilhelm, S., *et al.*, *Mutations towards enantioselectivity adversely affect secretion of Pseudomonas aeruginosa lipase*. FEMS microbiology letters, 2008. **282**(1): p. 65-72.
48. Eggeling, C., Berger, S., *et al.*, *Data registration and selective single-molecule analysis using multi-parameter fluorescence detection*. Jornal of Biotechnology, 2001. **86**(3): p. 163-80.
49. Kuhnemuth, R. and Seidel, C.A.M., *Principles of single molecule multiparameter fluorescence spectroscopy*. Single Molecules, 2001. **2**(4): p. 251-254.
50. Sindbert, S., Kalinin, S., *et al.*, *Accurate Distance Determination of Nucleic Acids via Forster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity*. Journal of the American Chemical Society, 2011. **133**(8): p. 2463-2480.
51. Karedla, N., Enderlein, J., *et al.*, *Absolute Photoluminescence Quantum Yield Measurement in a Complex Nanoscopic System with Multiple Overlapping States*. Journal of Physical Chemistry Letters, 2014. **5**(7): p. 1198-1202.
52. Case, D., Darden, T., *et al.*, *AMBER 11*. 2010, University of California: San Francisco.
53. Ryckaert, J.P., Ciccotti, G., *et al.*, *Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes*. Journal of Computational Physics, 1977. **23**(3): p. 327-341.

54. Darden, T., York, D., *et al.*, *Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems*. Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.
55. Roe, D.R. and Cheatham, T.E., *PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data*. Journal of Chemical Theory and Computation, 2013. **9**(7): p. 3084-3095.
56. Hummer, G. and Koefinger, J., *Bayesian ensemble refinement by replica simulations and reweighting*. Journal of Chemical Physics, 2015. **143**(24).
57. Boyd, S. and Vandenberghe, L., *Convex Optimization*. 2004, Cambridge: Cambridge University Press.

Supporting Information

Functional dynamics of a structurally minimalistic chaperone

Jakub Kubiak[1#], Neha Verma[2#], Holger Gohlke[2,4,5], Claus Seidel[1], Peter Dollinger[3],Filip Kovacic[3], Karl-Erich Jaeger[3,6]

[1]Institute of Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; [2]Heinrich-Heine-University of Düsseldorf, Institute of Pharmaceutical and Medicinal Chemistry, Düsseldorf, Germany; [3]Institute for Complex Systems - Structural Biochemistry (ICS-6), Forschungszentrum Jülich GmbH, Jülich, Germany; [4]John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC); [5]Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany; [6]Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

[#] These authors contributed equally to this work

Keywords*: *Pseudomonas aeruginosa*, molecular chaperone, steric chaperone, lipase, molecular dynamics simulations, FRET, protein folding, structural ensemble

Supplementary Figures

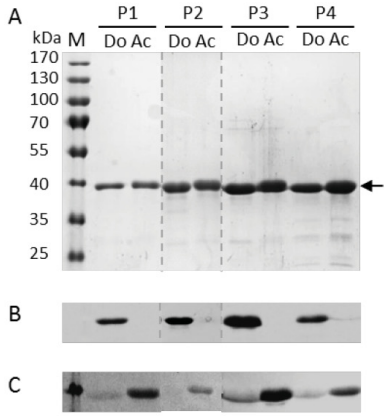


Figure S1: SDS-PAGE analysis of purification and fluorescence labeling of the Lif variants P1, P2, P3, and P4. Samples were analyzed on 12% SDS-PAGE and visualized with **A)** Coomassie Blue G-250 dye, **B)** illumination with ultra violet light and **C)** visible light. The numbers on the left-hand side indicate the molecular weight of standard proteins (M).

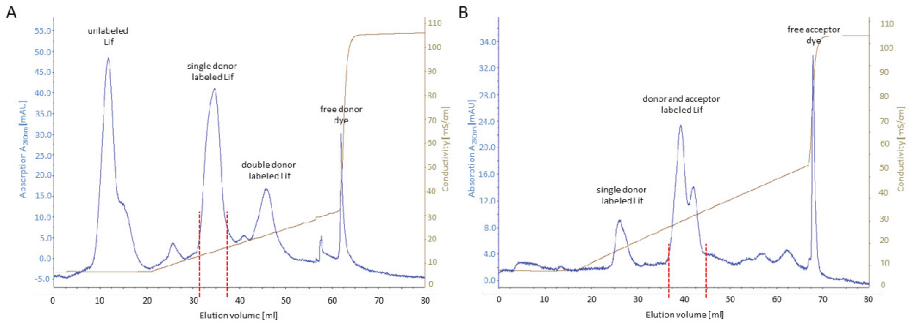


Figure S2: Ion exchange chromatographic purification of labeled Lif P1-2 variant after labelling with the donor (A) and the acceptor (B) dye. The blue diagram shows absorption profiles at 280nm. Brown diagram shows the ionic strength of the buffer. Collected fractions are indicated with red dotted lines.

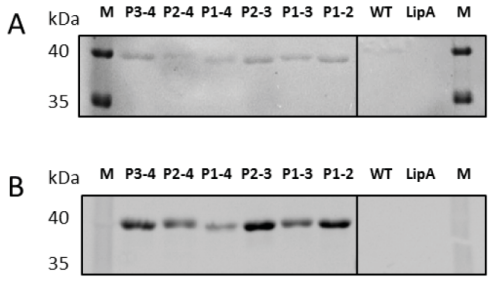


Figure S3: SDS-PAGE analysis of purification and fluorescence labeling of the Lif variants. Fluorescently labeled Lif variants including the wild type Lif (WT) and LipA were analyzed using 12% SDS-PAGE and visualized by illumination with visible (A), and ultra violet (B) light. The numbers on the left hand side indicate the molecular weight of standard proteins (M).

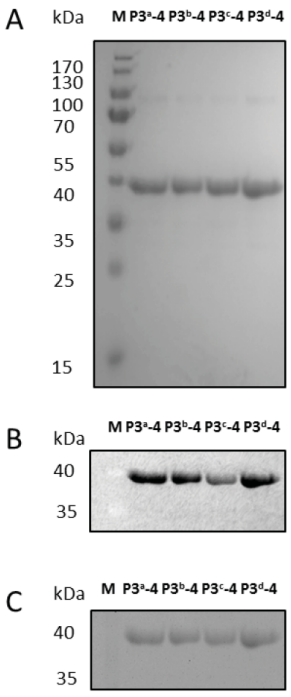


Figure S4: SDS-PAGE analysis of purification and fluorescence labeling of the Lif variants. Fluorescently labeled Lif variants including the wild type Lif (WT) and LipA were analyzed using 12% SDS-PAGE and visualized by staining with Coomassie Blue G-250 dye (A), illumination with visible (B), and ultra violet (C) light. The numbers on the left hand indicate the molecular weight of standard proteins (M).

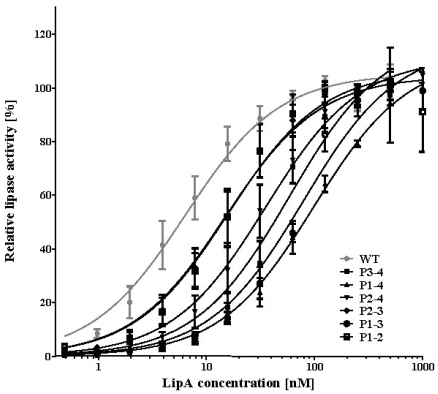


Figure S5: Kinetics of LipA activation with labelled Lif variants. Labeled Lif variants and the wild-type (WT) (2.5 nM) were titrated with pre-active LipA (from 0.5 to 1,000 nM), incubated over night at 4 °C in TB buffer followed by lipase activity assay. The activation constant (K_{act}) were calculated by one site binding hyperbola fit. Results are mean ± SEM of three independent experiments, with each set in triplicates.

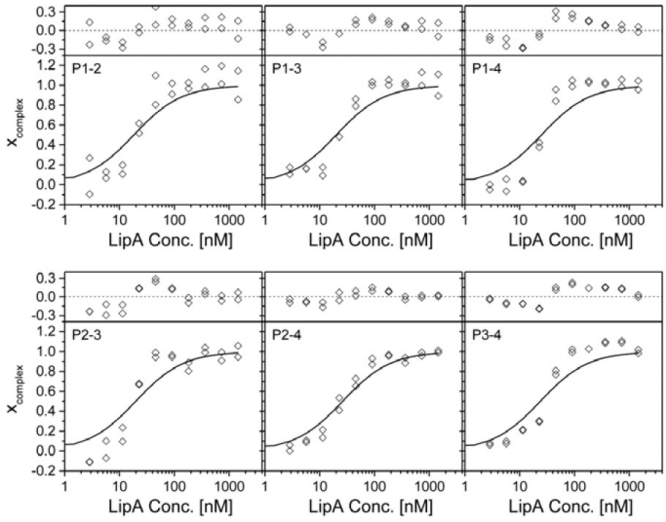


Figure S6: Equilibrium complex formation of labelled Lif variants and LipA. FRET variants of Lif labelled with Alexa488/647 were titrated with LipA. Proximity ratio $P = S_R/(S_G+S_R)$ was converted into fraction of the complex $x_{complex} = (P - P_0) / (P_{end} - P_0)$, where P_0 and P_{end} are proximity ratios of free Lif and Lif with maximum concentration of LipA. Binding model $x_{complex} = c_{LipA}/(K_D + c_{LipA})$, where K_D is equilibrium dissociation constant and c_{LipA} is concentration of LipA, was applied to 2 set of LipA dilutions

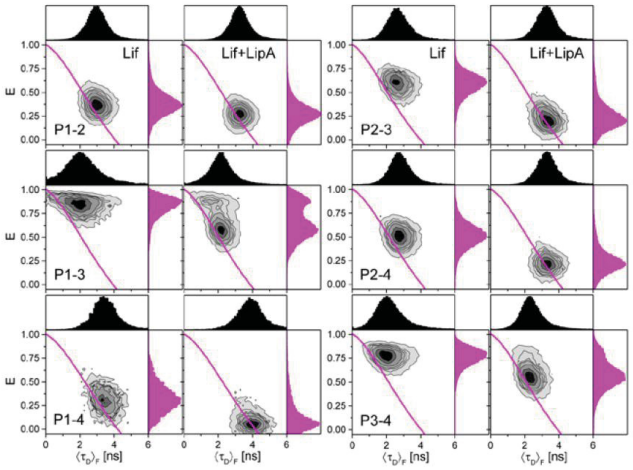


Figure S7: Single molecule MFD histograms. FRET efficiency E and burst-integrated fluorescence averaged lifetime of the donor $\langle\tau_D\rangle_F$ of six FRET variants of Lif labelled with Alexa488/647 are shown in absence and presence of a micromolar concentration of LipA. Linker-corrected static FRET lines (magenta), linking fluorescence observables assuming FRET distribution due to dye-linker diffusion of width (σ_{DA} = 6 Å), and parameters: green/red channel detection efficiency ratio γ' = 0.8, green and red channel background $\langle B_G\rangle$ = 0.55 kHz and $\langle B_R\rangle$ =0.35 kHz, direct excitation of acceptor by a blue laser = 0.02. Dye parameters in are collected in Table S4.

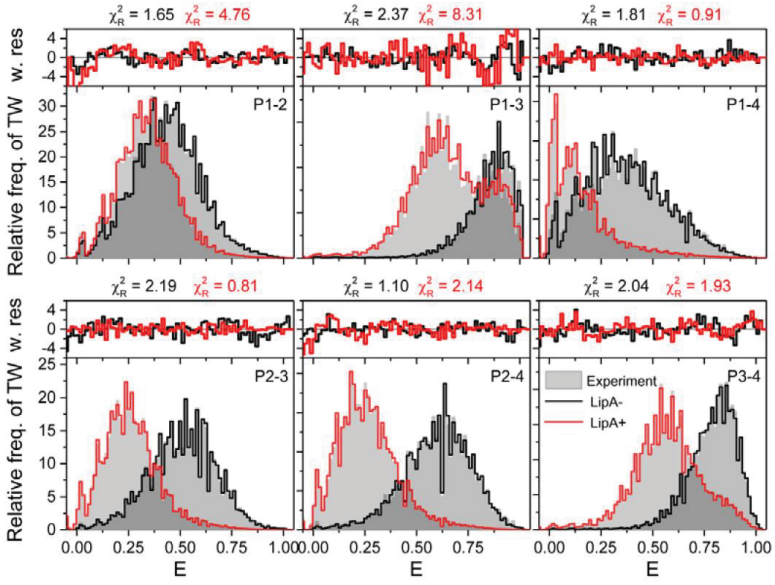


Figure S8: PDA of single molecule FRET efficiency histograms for labelled Lif variants in the absence (gray) and presence of a micromolar concentration of LipA (red). A single Gaussian

fit with variable fluorescence averaged distance $\langle R_{DA} \rangle_E$ and apparent width σ_{app} was applied. Observation window ΔTW = 1 ms, common fit parameters: green/red channel detection efficiency ratio γ' = 0.8, green and red channel background $\langle B_G \rangle$ = 0.55 kHz and $\langle B_R \rangle$ =0.35 kHz, direct excitation of acceptor by a blue laser = 0.02, Förster radius R_0 = 52 Å.

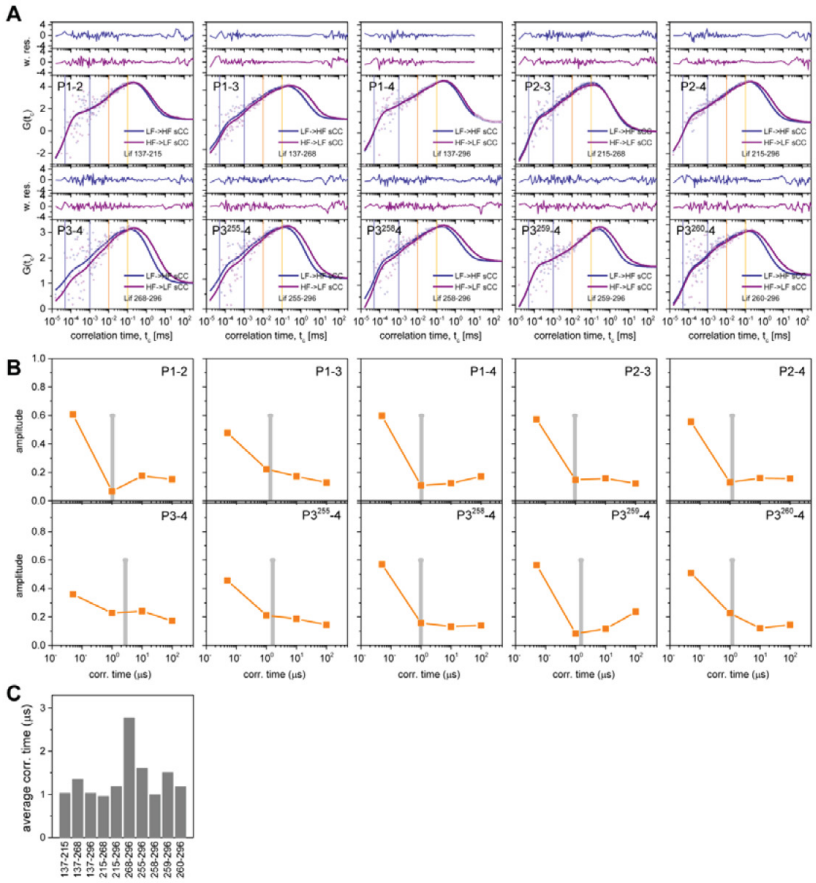
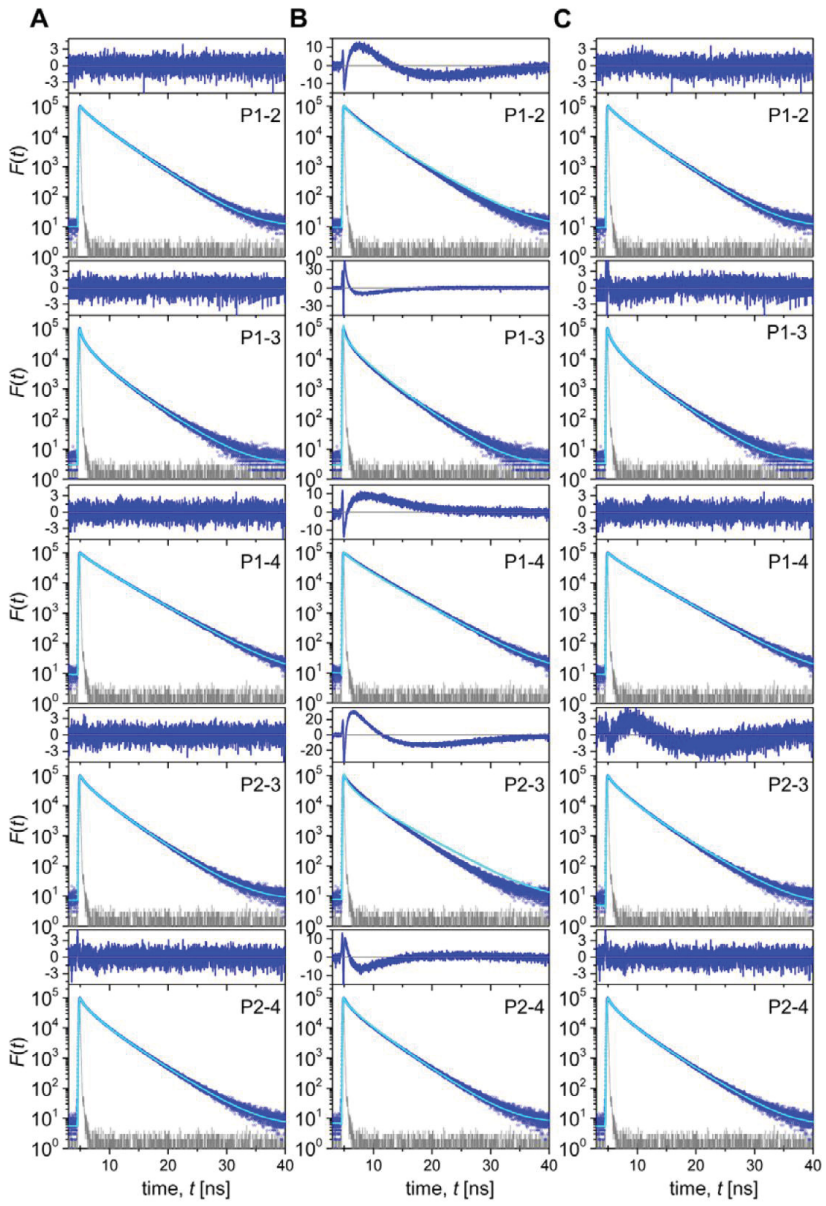


Figure S9: FRET dynamics probed by filter FCS. (**A**) Species cross-correlations (LF→HF and HF→LF) of labelled Lif variants created with HF- and LF-filters were fitted using a model with 3D Gaussian diffusion, a bunching term, asymmetric anti-correlation in millisecond time range indicating acceptor photo-bleaching, and 4 anti-correlation terms with relaxation times $t_{a,i}$ (50 ns, 1 µs, 10 µs, 100 µs), according to Eq.1 (see Table S7). (**B**) Kinetic spectra of FRET anti-correlation show a time-scale of exchange between FRET species. Log-averages are indicated by gray bars. (**C**) Log-average of anti-correlation time of Lif FRET variants shows FRET dynamics average in microsecond time-scale, with the exception of P3-4 variant (268-296) characterized by slower FRET dynamics.


A
B
C
P1-2
P1-3
P1-4
P2-3
P2-4
F(t)
time, t [ns]

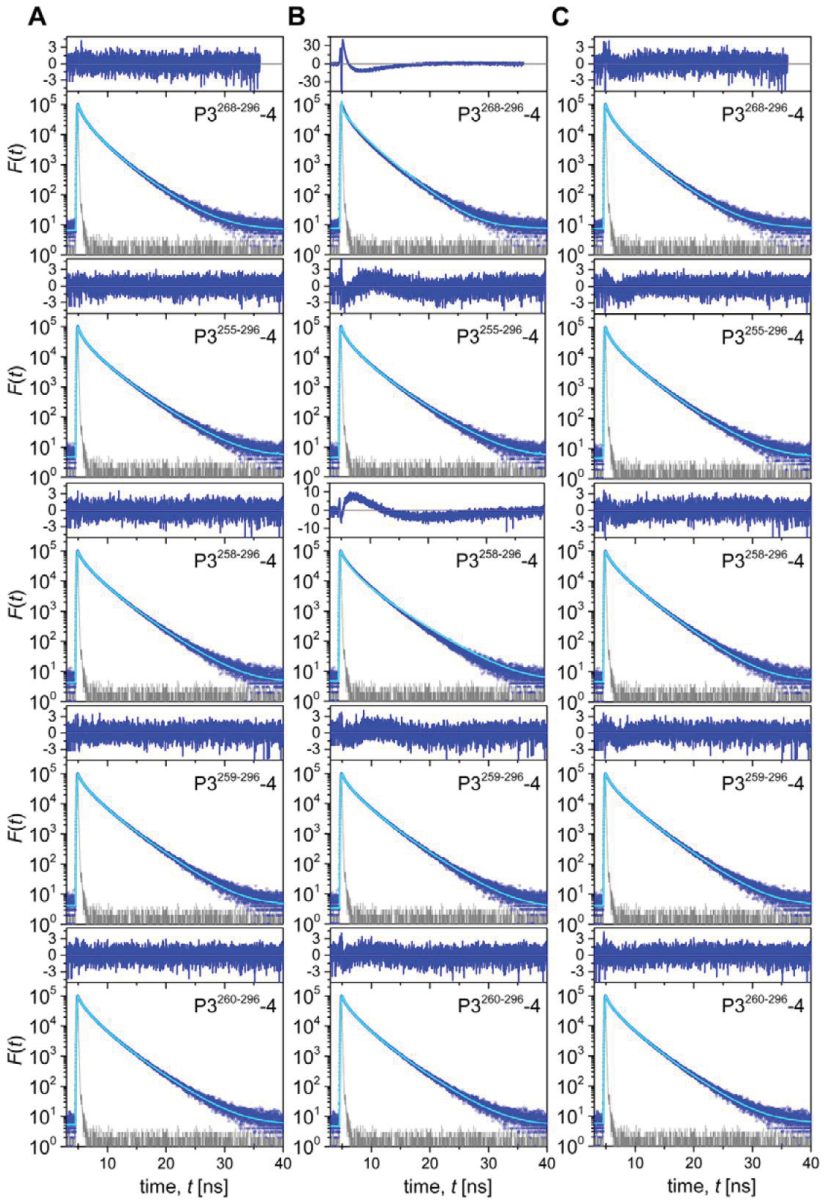


Figure S10: Time-resolved fluorescence intensity decay histograms (eTCSPC) of donor in presence of acceptor (donor reference not shown). Various models of $p(R_{DA})$ are fitted to the experimental decays: (A) generalized Gaussian distribution (Eq. 6), (B) initial MD ensemble, (C) MD ensemble weighted with $p(R_{DA})$ from generalized Gaussian distribution fit. Metrics of fit goodness (χ^2_R) are given in Table S8.

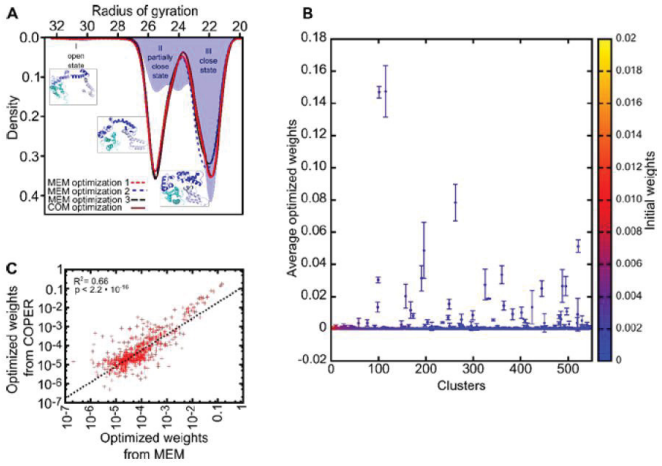


Figure S9: Reproducibility and validation of the maximum entropy method (MEM) used for cluster reweighting. **(A)** Frequency distribution of the radius of gyration (R_G) of the conformational ensemble of free Lif generated from ten unbiased MD simulations of 1 µs length (blue shaded region), showing four populations at ~21 Å, ~24 Å, ~25.7 Å, and one at ~31Å representing the closed, partially closed and open states of Lif. Exemplary conformations of the closed, partially closed and open state are shown as cartoon representation. The red, blue, and black curves show the frequency distribution after MEM optimizations and the brown curve after optimization by COPER. **(B)** Robustness of the MEM. The average optimized weight of each cluster ± SEM, computed across three independent optimization runs. Clusters are rank-ordered according to the initial weights shown by the color scale on the right. **(C)** Comparison of MEM with COPER. Optimized weights of clusters obtained by MEM are compared with optimized weights obtained by COPER at the entropy threshold.

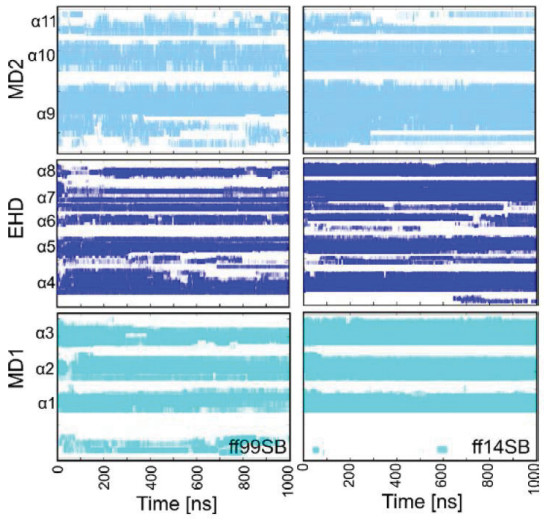


Figure S10. The α-helical content of separated domains of free Lif as a function of MD simulation time. The secondary structure of each domain residues MD1 (cyan, bottom), EHD

(light blue, middle), and MD2 (blue, top) are given as a function of the simulation time. In the left column, the secondary structure is shown for representative MD simulations performed with the ff99SB force field for all three domains and in right with ff14SB force field. All three helices in MD1 are stable over the simulation time irrespective of the used force field, whereas H5-H7, H9, and H11 show irregular α-helix formation irrespective of the force field.

Supplementary Tables

Table S1 Bacterial strains, plasmids, and oligonucleotides used in this work.

Name	Description	Reference/ source
Plasmids		
pEHTHis19	pET19b containing a 985 bp fragment carrying *lipH* gene with first 60 bp encoding the transmembrane domain replaced by a His_{10}-tag sequence from the plasmid, $PT7_{\phi 10}$, Ap^R	[28]
pETLipA-SS	pET22b containing a 916 bp fragment carrying lipA gene lacking its first 78 bp coding for the signal sequence, $PT7_{\phi 10}$, Ap^R	[28]
Strains		
E. coli BL21 (DE3)	F- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm* (DE3)	Novagen, Darmstadt, Germany
E. coli DH5α	*supE*44 Δ*lacU*169 (Φ80 lacZ ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1	Invitrogen, Karlsruhe, Germany
Oligonucleotides (5`→ 3`)[a]		
Q137C_for	CCAGGCGTTGGCGCTGA**T^{T}G^{G}T^C**GCCAATACATCGACTACAAGA	
Q137C_rev	TCTFTGTAGTCGATGTATTGGC**A^{A}C^{C}A^G**TCAGCGCCAACGCCTGG	
A215C_for	CGAGGAAAAGGCC**T^{G}G^{C}C^C**GCCATCGACCGC	
A215C_rev	GCGGTCGATGGC**G^{G}C^{G}A^C**GGCCTTTTCCTCG	
E268C_for[b]	GTCAGCAACTGGTGGGCGCCTGCGCCACCACCCGCCTGGAGC	
E268C_rev[b]	GGCGCCCACCAGTTGCTGACGC	
R296C_for	TTTCGCCGAGAAGAGC**T^{C}G^{G}C^G**ATCGAAGGCAATACCG	
R296C_rev	CGGTATTGCCTTCGAT**G^{C}C^{C}A^G**GCTCTTCTCGGCGAAA	
Lif_ΔVD_for[b]	AGGTTCCAGCCACATATGCTTGTCGTC	
Lif_ ΔVD_rev[b]	AGCATATGTGGCTGGAACCT GCCAAGGTCGCGCCGCTG	

[a]Mutated codons are indicated in bold, and nucleotides of the wild type gene are in superscript.
[b]Primers used for SLIC PCR.

Table S2 Activation (K_{act}) and equilibrium dissociation K_D constants for *in vitro* reaction between pre-active LipA and Lif.

	Lif variant	K_{act} ± SEM [nM]	K_D ± SEM [nM]
P1-2	Q137C-A215C	63 ± 24	18 ± 5
P1-3	Q137C-E268C	73 ± 9	20 ± 4
P1-4	Q137C-R296C	40 ± 6	25 ± 6
P2-3	A215C-E268C	43 ± 4	20 ± 5
P2-4	A215C-R296C	85 ± 9	26 ± 3
P3-4	E268C-R296C	24 ± 4	24 ± 5
WT		19 ± 2	N.D.*

* Dissociation constant for the wild type (WT) Lif could not be determined with FRET assay.

Table S3 Fluorescence parameters of single Cys Lif variants.

Lif variant		Dye	r_{steady}	$\langle\tau\rangle_X$ [ns]	$\langle\tau\rangle_F$ [ns]
P1	Q137C	Alexa488	0.15	3.8	4.1
P2	A215C	Alexa488	0.16	3.7	4.1
P3	E268C	Alexa488	0.24	3.8	4.0
P4	R296C	Alexa488	0.16	3.7	4.1
P1	Q137C	Alexa647	0.30	1.4	1.5
P2	A215C	Alexa647	0.29	1.5	1.6
P3	E268C	Alexa647	0.30	1.7	1.8
P4	R296C	Alexa647	0.27	1.6	1.7

Table S4 Average fluorescence quantum yields of donor and acceptor dye, $\Phi_{F,D}$ and $\Phi_{F,A}$, residual anisotropy r_{inf}, simulated κ^2 distribution with maximum error and maximum uncertainty of FRET distance estimation ΔR_{DA}.

Lif variant	$\Phi_{F,D}$	$\Phi_{F,A}$	$r_{inf,D}$	$r_{inf,A}$	$r_{inf,A(D)}$	κ^2 [a]	ΔR_{DA} [b]
P1-2	0.78	0.32	0.140	0.314	0.038	0.64±0.39	0.149
P1-3	0.74	0.34	0.151	0.309	0.022	0.64±0.39	0.158
P1-4	0.84	0.31	0.107	0.289	0.056	0.64±0.35	0.114
P2-3	0.77	0.38	0.199	0.322	0.039	0.72±0.40	0.143
P2-4	0.71	0.37	0.146	0.305	0.031	0.64±0.39	0.149
P3-4	0.76	0.36	0.195	0.282	0.032	0.65±0.40	0.158

[a] - Orientation factor κ^2 given as mean and standard deviation of the simulated distribution.
[b] - κ^2–based uncertainty of the donor-acceptor distance R_{DA}.

Table S5 Parameters of PDA Gaussian fit with variable fluorescence averaged distance $\langle R_{DA}\rangle_E$ and apparent width σ_{app}. Observation window ΔTW = 1 ms, common fit parameters: green/red channel detection efficiency ratio γ' = 0.8, green and red channel background $\langle B_G\rangle$ = 0.55 kHz and $\langle B_R\rangle$ =0.35 kHz, direct excitation of acceptor by a blue laser = 0.02, Förster radius R_0 = 52 Å,

Lif variant	LipA	$\langle R_{DA}\rangle_E$	σ_{app}	$x_{complex}$	χ^2_R	$\Phi_{F,A}$
P1-2	-	55	5		1.65	0.32
P1-2	+	59	3	0.87	4.76	
P1-3	-	38	5		2.37	0.34
P1-3	+	49	3	0.67	8.31	
P1-4	-	58	7		1.81	0.31
P1-4	+	77	7	0.83	0.91	
P2-3	-	51	4		2.19	0.38
P2-3	+	63	3	0.91	0.81	
P2-4	-	48	4		1.1	0.37
P2-4	+	63	5	0.96	2.14	
P3-4	-	41	4		2.04	0.36
P3-4	+	50	3	0.81	1.93	

Table S6 Fit parameters of fluorescence intensity donor decay (eq.4) and of FRET-induced donor decay with donor-acceptor distance distribution modeled by generalized Gaussian distribution (Eq. 6). Uncertainty denotes 0.99 confidence interval of the fit for given fitting parameter.

	P1-2	P1-3	P1-4	P2-3	P2-4	P3-4	P3^{255}-4	P3^{258}-4	P3^{259}-4	P3^{260}-4
χ^2_R	1.060	1.073	1.080	1.073	1.103	1.103	1.086	1.115	1.085	1.064
					$p(R_{DA})$ model parameters, eq. 6					
$\langle R_{DA} \rangle$, Å	65 ±1	29 ±7	71 ±2	59 ±1	53 ±2	46 ±1	46 ±1	52 ±1	53 ±1	53 ±1
σ_{DA}, Å	9.3 ±1.9	34 ±14	15 ±2.1	12 ±1.6	16 ±1.7	13 ±0.5	16 ±1	12 ±0.5	13 ±0.5	11 ±0.5
κ	0.28 ±0.19	0.53 ±0.32	0.11 ±0.21	0.41 ±0.2	-0.07 ±0.21	0.34 ±0.12	0.15 ±0.16	0.37 ±0.14	0.24 ±0.13	0.36 ±0.14
x_{D0}	0.030	0.004	0.410	0.043	0.023	0.014	0.018	0.022	0.008	0.031
					Donor parameters $F_{D(0)}(t)$, eq. 4					
τ_1, ns	4.23	4.23	4.20	4.19	4.18	4.23	4.16	4.22	4.16	4.15
τ_2, ns	2.28	2.29	2.30	2.16	2.12	2.23	2.21	2.17	2.16	2.15
τ_3, ns	0.376	0.385	0.382	0.335	0.357	0.385	0.417	0.405	0.413	0.429
x_1	0.805	0.806	0.819	0.824	0.768	0.806	0.701	0.740	0.688	0.702
x_2	0.116	0.116	0.109	0.102	0.133	0.116	0.180	0.156	0.187	0.182
x_3	0.079	0.078	0.072	0.074	0.099	0.078	0.119	0.104	0.125	0.116

Table S7 (A) Parameters of filtered FCS fit (Eq. 1). LF-HF denotes low-FRET to high-FRET species cross-correlation, and HF-LF the reverse. Single field for 2 cross-correlation denotes jointly fitted parameter.

	P1-2		P1-3		P1-4		P2-3		P2-4	
	LF-HF	HF-LF	LF-HF	HF-LF	LF-HF	HF-LF	LF-HF	HF-LF	LF-HF	HF-LF
χ^2	0.613	0.575	0.472	0.543	0.357	0.336	0.575	0.540	0.530	0.382
					G_{dif} (3D Gaussian diffusion)					
N	0.255	0.097	0.747	0.175	0.271	0.216	0.185	0.117	0.219	0.053
t_{dif}, ms	2.4		2.3		2.5		1.5		1.7	
ω_0/z_0					5					
					G_b (bunching)					
b	0.539	0.647	0.545	0.681	4.004	3.423	0.248	0.290	0.331	0.649
t_b, ms	3.5	3.5	5.1	5.1	24.7	24.7	3.0	3.0	4.1	4.1
					G_{ph} (acceptor photo-bleaching, only in HF-LF correlation)					
a_{ph}	0.000	0.634	0.000	0.781	0.000	0.251	0.000	0.390	0.000	0.770
t_{ph}, ms					10					
					G_a (anti-correlation)					
a	2.27	2.28	2.03	2.17	2.38	2.41	1.77	1.83	1.89	1.92
a_1	0.151		0.155		0.183		0.122		0.156	
a_2	0.176		0.154		0.116		0.158		0.160	
a_3	0.067		0.224		0.111		0.148		0.130	
a_4	0.606		0.466		0.589		0.572		0.554	
t_{a1}, µs					100					
t_{a2}, µs					10					
t_{a3}, µs					1					
t_{a4}, ns					50					

(B) Parameters of filtered FCS fit (Eq. 1). LF-HF denotes low-FRET to high-FRET species cross-correlation, and HF-LF the reverse. Single field for 2 cross-correlation denotes jointly fitted parameter.

	P3-4		P3^{255}-4		P3^{258}-4		P3^{259}-4		P3^{260}-4	
	LF-HF	HF-LF	LF-HF	HF-LF	LF-HF	HF-LF	LF-HF	HF-LF	LF-HF	HF-LF
χ^2	0.596	0.495	0.494	0.789	0.598	0.744	0.713	0.569	0.788	0.662
G_{dif} (3D Gaussian diffusion)										
N	0.411	0.100	0.512	0.142	0.582	0.247	0.416	0.123	0.324	0.078
t_{dif}, ms	2.1		2.2		2.3		2.4		1.6	
ω_0/z_0	5									
G_b (bunching)										
b	0.473	0.562	0.513	0.595	0.611	0.395	0.559	0.642	0.360	0.521
t_b, ms	4.5	4.5	4.0	4.0	6.3	6.3	4.0	4.0	4.0	4.0
G_{ph} (acceptor photo-bleaching, only in HF-LF correlation)										
a_{ph}	0.000	0.763	0.000	0.726	0.000	0.573	0.000	0.707	0.000	0.769
t_{ph}, ms	10									
G_a (anti-correlation)										
a	1.23	1.44	1.40	1.59	2.29	2.54	2.22	2.21	1.80	1.87
a_1	0.173		0.145		0.141		0.236		0.145	
a_2	0.241		0.187		0.132		0.116		0.121	
a_3	0.227		0.212		0.158		0.084		0.226	
a_4	**0.359**		0.456		0.569		0.564		0.508	
t_{a1}, µs	100									
t_{a2}, µs	10									
t_{a3}, µs	1									
t_{a4}, ns	50									

Table S8 Goodness of fit χ^2_R of FRET-induced donor decay to (A) generalized Gaussian distribution (Eq. 6), (B) $p(R_{DA})$ from initial MD ensemble, (C) $p(R_{DA})$ from MD ensemble weighted with $p(R_{DA})$ from generalized Gaussian distribution fit.

	χ^2_R		
	A	C	B
P1-2	1.060	21.98	1.136
P1-3	1.073	36.95	1.320
P1-4	1.080	14.29	1.082
P2-3	1.073	114.6	2.579
P2-4	1.103	8.208	1.108
P3-4	1.103	48.44	1.299
P3^{255}-4	1.086	1.417	1.173
P3^{258}-4	1.115	8.526	1.242
P3^{259}-4	1.085	1.228	1.152
P3^{260}-4	1.064	1.151	1.114

11. ORIGINAL PUBLICATION II (Submitted 2019)

Structural and dynamic insights revealing how lipase binding domain MD1 of *Pseudomonas aeruginosa* foldase affects lipase activation

Aldino Viegas[1,$,#], Peter Dollinger[2,$], Neha Verma[3,$], Jakub Kubiak[5], Thibault Vienne[1,4], Claus A.M. Seidel[5], Holger Gohlke[3,4,6,*], Manuel Etzkorn[1,4,7*], Filip Kovacic[2] and Karl-Erich Jaeger[2,8,*]

Running title: **Structural insights into foldase-mediated lipase activation**

From the [1]Institute of Physical Biology, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; [2]Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany; [3]Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; [4]Institute for Complex Systems ICS-6: Structural Biochemistry, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany; [5]Institute of Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; [6]John von Neumann Institute for Computing (NIC) and Jülich Supercomputing Centre (JSC), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany; [7]JuStruct: Jülich Center for Structural Biology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany; [8]Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

[#]Present address: Departamento de Química, Faculdade de Ciências e Tecnologia UCIBIO-REQUIMTE, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

*To whom correspondence should be addressed: Karl-Erich Jaeger: Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, D-52425, Jülich, Wilhelm Johnen Str. 1, Karl-Erich.jaeger@fz-juelich.de; Tel. (+49)02461616966; Fax. (+49)02461612490. Manuel Etzkorn: Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, D-40225, Düsseldorf, Universitätsstr. 1; manuel.etzkorn@hhu.de; Tel. (+49)2118112023; Fax. (+49)2118113847. Holger Gohlke: Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, D-40225, Düsseldorf, Universitätsstr. 1; gohlke@uni-duesseldorf.de; Tel. (+49) 2118113662; Fax. (+49)2118113847.

[$]These authors contributed equally to this work.

Keywords: *Pseudomonas aeruginosa*, lipase, chaperone, protein folding, virulence factor, solution NMR, long-range network interactions, constraint network analysis

---

ABSTRACT

Folding and cellular localization of many proteins of Gram-negative bacteria rely on a network of chaperones and secretion systems. Among them is the lipase-specific foldase Lif, a membrane-bound steric chaperone that tightly binds (K_D = 29 nM) and mediates folding of the lipase LipA, a virulence factor of the pathogenic bacterium *P. aeruginosa*. Lif consists of five-domains, including a mini domain MD1 essential for LipA folding. However, the molecular mechanism of Lif-assisted LipA folding remains elusive. Here, we show in *in vitro* experiments using a soluble form of Lif (*s*Lif) that isolated MD1 inhibits *s*Lif-assisted LipA activation. Furthermore, the ability to activate LipA is lost in the variant *s*Lif$_{Y99A}$, in which the evolutionary conserved amino acid Y99 from helix α1 of MD1 is mutated to alanine. This coincides with an approximately three-fold reduced affinity of the variant to LipA together with increased flexibility of *s*Lif$_{Y99A}$ in the complex as determined by polarization-resolved fluorescence spectroscopy. We have solved the NMR solution structures of *P. aeruginosa* MD1 and variant MD1$_{Y99A}$ revealing a similar fold indicating that a structural modification is likely not the reason for the impaired activity of variant *s*Lif$_{Y99A}$. Molecular dynamics simulations of the *s*Lif:LipA complex in connection with rigidity analyses suggest a long-range network of interactions spanning from Y99 of *s*Lif to the active site of LipA, which might be essential for LipA activation. These findings provide important details about the putative mechanism for LipA activation and point to a general

mechanism of protein folding by multi-domain steric chaperones.

The Gram-negative human pathogen *Pseudomonas aeruginosa* produces a wide range of extracellular enzymes,(1, 2) among them the lipase LipA, a secreted putative virulence factor(3-5). For its conversion into an enzymatically active conformation, LipA requires the assistance of an inner membrane-bound chaperone named lipase-specific foldase (Lif)(6). On the folding pathway, LipA can adopt several structurally different intermediates: an inactive and unfolded molten globule-like conformation(7), a near-natively folded pre-active conformation(8) and two folded conformations that differ in the structure of the α-helical lid covering the active site, with the folded closed conformation being enzymatically inactive and the folded open conformation enzymatically active(9, 10). Addition of Lif to the pre-active lipase immediately activates the folding intermediate(11-16), suggesting that the interactions with Lif help overcoming an energetic barrier on the folding pathway of lipase LipA.

Lif proteins constitute a unique class of steric chaperones(17, 18). *P. aeruginosa* Lif has a five-domain organization with a transmembrane α-helical domain (TMD), followed by a probably unstructured variable linker domain (VLD) and the catalytic folding domain (CFD) which interacts with the lipase (Fig. 1). The crystal structure of *Burkholderia glumae* foldase (homologous to *P. aeruginosa* foldase) in complex with its cognate lipase reveals only the periplasmic catalytic folding domain(10). This domain consists of 11 α-helices connected by loops and is organized into two globular domains, mini-domain 1 (MD1, α1-α3) and mini-domain 2 (MD2, α9-α11), which are connected by the highly flexible extended helical domain (EHD, α4-α8). Six α-helices of Lif (α1, α4, α5, α7, α9, α11) are in direct contact with LipA, forming a notably large interface between Lif and LipA, which is consistent with the high binding affinity in the nanomolar range of these two molecules(10).

The sequence alignment of *P. aeruginosa* foldase (*Pa*Lif) and *B. glumae* foldase (*Bg*Lif), the only foldase with known 3D structure, revealed that among the five domains MD1 shares the highest sequence similarity (52%) (Fig. S1). Similar sequences often exert similar functions, which holds for MD1, as the chimeric foldase of *B. glumae* containing the MD1 of *P. aeruginosa* Lif still activated *B. glumae* LipA(8). In contrast, other hybrid *B. glumae*-*P. aeruginosa* Lifs with replaced MD2 and EHD were inactive and *B. glumae* Lif did neither activate *P. aeruginosa* LipA nor *vice versa*(8). The importance of MD1 for foldase activity was further highlighted by the finding that MD1 comprises the foldase sequence motif RXXFDY(F/C)L(S/T)A (X can be any residue) which is evolutionarily strongly conserved among all foldase families(12) and which when mutated inactivates foldase(19). However, despite this detailed knowledge, the molecular mechanism of foldase-assisted lipase folding still remains elusive.

Here, we investigated the role of MD1 for the activation of pre-active LipA by biochemical analysis, NMR spectroscopy, fluorescence spectroscopy and molecular simulations. Our solution NMR structures reveal that mutation Y99A in MD1 induces only slight changes in the protein structure. However, our biochemical activation assays show that in contrast to MD1, $MD1_{Y99A}$ does not decelerate *s*Lif-induced activation of LipA. sLif is a soluble form of *Pa*Lif that lacks the TMD and Y99 is evolutionary conserved and located in helix α1 of MD1. While sLif and $sLif_{Y99A}$ both interact with LipA, fluorescence-based assays demonstrate that the mutation significantly reduces the sLif-LipA affinity. The role of mutation Y99A on LipA activation was probed by molecular dynamics (MD) simulations and rigidity theory. Comparative constraint network analyses of MD-generated conformational ensembles of wild-type *s*Lif and variant $sLif_{Y99A}$ in complex with LipA suggest that long-range network interactions spanning from Y99 of *s*Lif to the active site of LipA are likely involved in LipA activation.

Results

Isolated MD1, but not $MD1_{Y99A}$ decelerates sLif-induced activation of LipA

In line with previous results(19) we observed that the point mutation generating variant $sLif_{Y99A}$ strongly modifies *in vitro* folding of LipA (*s*Lif lacks the TMD, which is dispensable for *in vitro* Lif function (14)) (Fig. 2A). Interestingly, however, LipA strongly binds to both *s*Lif and variant $sLif_{Y99A}$ (Fig. 2B).

Presumably, specific interactions of MD1 with LipA are important for its activation as proposed for activation of *B. glumae* lipase, too.(8, 10, 19) We purified MD1 (Fig. S2) and demonstrated that these interactions are not sufficient for LipA activation as isolated MD1 could not activate pre-active LipA *in vitro* (data not shown). However, the addition of MD1 to pre-active LipA in 12 to 20-fold molar excess during activation of LipA with *s*Lif significantly ($p < 0.001$, $n = 4$) slowed down the activation reaction (Fig. 2C). This result indicates that isolated MD1 can interfere with *s*Lif's capability to activate LipA. This effect was not observed with isolated $MD1_{Y99A}$ (Fig. 2C). We analyzed the affinity of both *s*Lif and variant *s*Lif_{Y99A} to LipA using a fluorescence-based assay (Tab. S1 and S2). We observed stronger binding of *s*Lif (K_D = 29 ± 9 nM) than *s*Lif_{Y99A} (K_D = 77 ± 24 nM) (Fig. 2D).

LipA complexes with s*Lif and variant* sLif_{Y99A} *exhibit similar unfolding profiles*

We further probed the interactions of MD1, $MD1_{Y99A}$, *s*Lif and *s*Lif_{Y99A} with pre-active LipA by analyzing the intrinsic protein fluorescence of respective solutions during thermal unfolding using a Prometheus nanoDSF device. While *s*Lif and *s*Lif_{Y99A} alone show typical unfolding curves with unfolding temperatures of ~50°C (Fig. 3A), the thermal unfolding of MD1 and $MD1_{Y99A}$ cannot be monitored with nanoDSF because MD1 does not contain tryptophan residues. The thermal unfolding curve of pre-active LipA does not show a sharp unfolding transition typical for folded proteins but rather a broad transition with a maximum at 73°C. Addition of MD1 or $MD1_{Y99A}$ to pre-active LipA did not considerably affect this unfolding curve, indicating that the domains do not change the fold of pre-active LipA and/or do not strongly interact with LipA (Fig. S3), which was also confirmed by fluorescence binding assay (Figs. S4 and S5 and Tab. S3).

In contrast, the addition of *s*Lif or *s*Lif_{Y99A} to pre-active LipA yielded typical unfolding curves with a maximum of the first derivatives at ~37°C (Fig. 3A). Further heating of the solutions above 37°C resulted in unfolding curves similar to the one observed for pre-active LipA (Fig. 3A). The unfolding temperatures for *s*Lif or *s*Lif_{Y99A} in the presence of LipA are ~7°C higher than the temperature at which LipA activity is reduced to 50% in a temperature-dependent lipase activity assay in the presence of *s*Lif (half-inactivation temperature, T_{50} = 29.0 ± 0.2°C, Fig. 3B); the temperature difference may be explained by the perturbation of the contacts within *s*Lif:LipA complex.

Internal flexibility of variant sLif_{Y99A} *in complex with LipA is increased as compared to* s*Lif*

Rotational diffusion and flexibility of *s*Lif were investigated by different techniques of polarization-resolved fluorescence spectroscopy.(20) The fluorescence anisotropy decay $r(t_c)$ of the fluorescent probe, Bodipy FL NHS ester (BDP FL) conjugated to *s*Lif or *s*Lif_{Y99A}, is sensitive to local flexibility (time scale of 0.1 to 10 ns) and global rotation of the labelled molecules (Figs. 3C and S6). Typical fluorescence lifetime data and depolarization times are compiled in Tab. S4. Polarization-resolved full-FCS (pFCS) is sensitive to depolarization motions in a time range > 10 ns (Fig. 3D), so that a joint analysis of time-resolved anisotropy and pFCS can capture a much wider time range (0.1 ns to ms) (Fig. 3E).

Considering free sLif and sLif_{Y99A}, the fluorescence anisotropy (Fig. 3C) and pFCS (Fig. 3D) show that *s*Lif and *s*Lif_{Y99A} exhibit similar hydrodynamic properties (rotational correlation time global ρ_{global} = 33 ns). Similar high amplitudes of pico- to nanosecond dynamics are evidence for a high degree of internal flexibility of the protein in the absence of LipA (order parameter, S^2 = 0.30, Fig. 3C, Tab. S4) that usually results in dynamic conformational ensemble,(21) which agrees with the findings of (7).

For the complexes *s*Lif:LipA and *s*Lif_{Y99A}:LipA the global rotational correlation time global is nearly doubled (ρ_{global} = 50 ns). This experimental value agrees with the structure-based theoretical value for the global rotational correlation time ρ_{global} = 45 ns obtained by HydroPRO(22) using the crystal structure of the complex of *B. glumae* PDB code 2ES4.(10) The fact that ρ_{global} in free *s*Lif is reduced indicates that its conformation must be more collapsed. The comparison of order parameters in Fig. 3E gives further insights into the internal interactions. Notably, the comparison shows that the two independent fluorescence techniques agree very well with respect to observed depolarization times and amplitudes.

In the case of the *s*Lif:LipA complex, the increased order parameter S^2 = 0.38 for global motion (blue arrow in Fig. 3E) indicates that *s*Lif is less flexible in its bound form. However, for

*s*Lif$_{Y99A}$:LipA, the order parameter for global motion (S^2 = 0.22, orange arrow) is even further decreased as compared to free *s*Lif$_{Y99A}$ (S^2 = 0.30, Fig. 3E). This finding agrees well with the result that *s*Lif$_{Y99A}$:LipA forms a less tight complex (Fig. 2D) as compared to wt *s*Lif. At the same time, the complex formation of *s*Lif with LipA disrupts many internal interactions in free *s*Lif so that internal friction should be reduced due to missing contacts. To conclude, the number of contacts in the *s*Lif$_{Y99A}$:LipA complex is significantly reduced. As the Y99A mutation is in MD1 domain, we can further conclude that for the mutant the MD1 domain does not, or only weakly, interact with LipA while the rest of *s*Lif should still interact normally.

Structural insights into MD1s of P. aeruginosa *Lif and variant Lif$_{Y99A}$*

So far, a high-resolution structure of Lif from *P. aeruginosa* does not exist; this is also true for each of the individual Lif domains. To obtain the first structural insights into this system and to investigate the effects of the critical Y99A mutation on the structure of the MD1 domain, we here solved the NMR solution-structure of the isolated MD1 domain (Fig. 4A, PDB code 5OVM; BMRB code 34175) as well as of the MD1$_{Y99A}$ variant (Fig. 4B, PDB code 6GSF; BMRB code 34286).

Both MD1 and MD1$_{Y99A}$ resemble a three α-helical bundle preceded by 27 N-terminal residues without clear secondary structure. Only minor structural differences were observed within each ensemble of 20 energetically most favorable structures for MD1 as well as for MD1$_{Y99A}$, as indicated by RMSD$_{Cα}$ of 1.3 ± 0.3 Å and 0.8 ± 0.2 Å, respectively. The obtained structures of the isolated MD1 variants are similar to the respective domain in the crystal structure of the Lif:LipA complex from *B. glumae* (Fig. 4C),(10) showing that this domain adopts a stable fold, even when isolated and in the absence of a lipase.

Overall, both variants from *P. aeruginosa* exhibit rather similar 3D structures, with an RMSD$_{Cα}$ of 2.4 Å, when comparing the MD1 and MD1$_{Y99A}$ structural ensembles. This shows that the Y99A mutation does not alter the overall fold of MD1. Nevertheless, some differences are still visible when comparing both structures. These differences include (i) helix 2, which is slightly tilted in the MD1$_{Y99A}$ variant as compared to MD1, as well as (ii) the degree of 'disorder' of the N-terminal coil including the loop interacting with helix 1. Yet, the second difference may be a direct consequence of limited distance restraints due to chemical exchange of the amide protons in this part of the protein.

A comparison of the ^{1}H-^{15}N HSQC spectra for both MD1 variants reveals a rather high amount of chemical-shift perturbations induced by the point mutation (Figs. 5A, B). Mapping the strongest perturbations on the 3D structure suggests that the mutation does not only affect the chemical environment of its direct neighbors but does induce effects in several areas of the protein (Fig. 5C). This observation is consistent with the differences found in the structures for both variants, in particular the relative tilt of helix 2 in the protein core. Yet, when comparing the ^{13}C chemical shifts ($C_α$, $C_β$), which are particularly sensitive to the secondary structure, only minor differences are found (Fig. 5D), again in accordance with the observed structural differences.

In agreement with the previously determined crystal structure of the homologous domain from *B. glumae*,(10) our data show that the N-terminal part does not exhibit a clear secondary structure motif. Furthermore, for larger parts of this region, the amide proton could not be detected, in line with elevated chemical exchange, indicating the absence of a hydrogen-bond network and/or the absence of protective steric effects provided by the remaining residues of the three-helix bundle.

However, our data also show that the N-terminal region is not completely disordered. On the one hand, this is confirmed by the rather low mobility seen in NMR dynamics data for several residues with observable amide protons (Figs. S7A-C). In fact, the loop connecting helix 1 and 2 appears to be more flexible than the N-terminal segment. On the other hand, secondary chemical shift analysis of ^{13}C frequencies, which could be assigned for most N-terminal residues, differs substantially from a pure random coil character (Fig. 5D). Furthermore, clear inter-residue NOE correlations connect the N-terminal region to the three-helix bundle (Fig. 5E). Within the here detected parameters the features of the N-terminal region are similar in both MD1 variants.

To gain further insights into the effect of the Y99A mutation on the interaction of MD1 with LipA we acquired ^{1}H,^{15}N-HSQC spectra of MD1 and MD1$_{Y99A}$ in the presence and absence

of three-fold molar excess of LipA (Fig. S8). The data reveals a clearly noticeable signal decrease induced by the presence of LipA for MD1. In general, such a signal decrease can either indicate the formation of a tightly bound complex, which, however, is too large to be detected in the NMR spectrum. Therefore, only the reduced signal of the unbound state is detected (so-called NMR slow exchange regime). Based on fluorescence measurements obtained under comparable conditions it can be excluded that a tight complex is formed between MD1 and LipA (Fig. S5). Alternatively, the signal decrease can be explained by a transient interaction with exchange rates in the range of the NMR timescale, leading to peak broadening (so-called NMR intermediate exchange regime). Interestingly, the observed signal decrease is considerably stronger for MD1 as compared to $MD1_{Y99A}$ suggesting that the mutation further reduces the interaction between MD1 and LipA. This observation is in line with the different behavior found in our inhibition assay (Fig. 2C) as well as the reduced LipA affinities observed for *s*Lif and $sLif_{Y99A}$ (Fig. 2D). Albeit the NMR data reveal minor but noticeable structural differences in MD1 structure due to the Y99A mutation and potential modulations in LipA affinity, it is at this point unclear how these differences can modulate Lif's capability to fold LipA.

Lif_{Y99A} exerts a long-range effect on LipA which may destabilize the structure of the substrate-binding pocket.

To understand the possible role of the mutation Y99A for Lif-induced activation of LipA, we initially compared the X-ray structure of *P. aeruginosa* LipA in the open, active conformation (PDB code 1EX9) with the X-ray structures of *B. glumae* lipases in their closed, inactive conformations, one in the complex with its specific foldase (PDB code 2ES4) and the other in the unbound form (PDB code 1QGE) (Figs. 6A-C). In the open conformation, helix α5 is moved away from the active site, allowing substrates to access this site and a short two-stranded β-sheet close to the active site is formed by residues 21-22 and 25-26. The enclosing residues 17-30 shape the substrate-binding pocket (*23*) (Fig. 6D). Because this β-sheet is not formed in the closed, unbound conformation of LipA, where helix α5 is occluding the active site, we postulated this structural element as a hallmark of the open and active LipA. Notably, the foldase-bound lipase, with an overall and active site structure mainly identical to those of unbound lipase(9, 10) does have a two-stranded β-sheet in the region of residues 17-30, yet helix α5 is in the closed conformation. The structural comparison thus indicates that the foldase induces the formation of the two-stranded β-sheet during activation of the lipase. Hence, the foldase-bound lipase can apparently be considered an intermediate conformation on its way to an open conformation, with residues 17-30 acting as a "loaded spring".

The effect of mutation Y99A in Lif for lipase activation was further examined. We first generated a homology model of the *P. aeruginosa* *s*Lif:LipA complex, since no experimental structure of *P. aeruginosa* Lif has been reported so far. As a template structure, we used the structure of the *B. glumae* foldase-lipase complex (PDB code 2ES4) (see Materials and Methods section for details). The final model was assessed with our in-house model quality assessment program TopScore(24, 25) and found to be 68 % correct for LipA and 52 % correct for Lif (Fig. 6E). The final *s*Lif:LipA model was used as an input structure to perform ten independent all-atom MD simulations of 1.5 μs length each.

To investigate the influence of the Y99A mutation on the structural stability of LipA, we used an ensemble-based perturbation approach(26) integrated into the CNA approach, a method for analyzing biomolecular rigidity and flexibility.(26) CNA was applied on the conformational ensemble of the *s*Lif:LipA complex generated from the above MD simulations, constituting the ground state (see Materials and Methods section for details). A perturbed state of the *s*Lif:LipA complex was generated by substituting Y99 with alanine, but keeping the structures of *s*Lif and LipA unchanged otherwise. We followed that approach because LipA strongly binds to both *s*Lif variants (Fig. 2B), suggesting that the respective complex structures are very similar in both cases. The computed changes in the residue-wise free energy $\Delta G_{i,\mathrm{CNA}}$ (eq. 1), a measure for structural stability,(*26*) was largest for residues 1-45, 197-202, 242-250 and 268-286 in LipA (Figs. 6F, G).

Notably, these affected residues form a narrow pathway that reaches the β-sheet-forming region including residues 28-30 (Figs. 6F, G), indicating that the mutation decreases the

stability of substrate binding pocket (Fig. 6D). We speculate that this decrease in stability prevents the partial β-sheet formation of substrate binding pocket and, thus, disfavors LipA activation by sLif$_{Y99A}$. Finally, the Y99A mutation also affects the stability of a number of residues in sLif itself, especially the neighboring residues 66-80 and 89-115 in MD1 (Fig. 6H), in agreement with the above NMR data (Fig. 5A). In summary, our CNA results indicated that the Y99A mutation in MD1 of sLif exerts a long-range influence on LipA structural stability that reaches the substrate-binding region that forms a β-sheet upon activation of LipA.

Discussion

In this work, we studied the role of MD1 and of MD1's residue Y99 on the activation of pre-active LipA by biochemical analysis, NMR and fluorescence spectroscopy, as well as molecular simulations. We show that the Y99A mutation in MD1 induces only minor changes in the domain's overall structure. However, the activation of LipA induced by sLif is inhibited by addition of MD1, which is not seen when using variant MD1$_{Y99A}$ (Fig. 2C). Comparative CNA suggested that long-range network interactions span from Y99 of sLif to the active site of LipA that are likely involved in LipA activation (Fig. 6).

MD1 was suggested as an essential domain for Lif activity as it contains an amino acid sequence motif conserved among all foldases,(12) which upon mutation leads to Lif inactivation.(19) We confirmed this finding by showing that the *P. aeruginosa* sLif$_{Y99A}$ variant carrying a single amino acid mutation in the foldase motif did not activate *P. aeruginosa* LipA, in contrast to sLif (Fig. 2A).

Previously, Shibata and coworkers reported that *P. aeruginosa* Lif$_{Y99C}$, Lif$_{Y99H}$, Lif$_{S102R}$ and Lif$_{R115C}$ variants (all carrying mutations in the foldase motif) do not form complexes with LipA.(19) By contrast, we showed, using co-purification and fluorescence-based assays, that sLif$_{Y99A}$ binds pre-active LipA. Both proteins were co-purified by affinity chromatography and sLif$_{Y99A}$ in the presence of LipA yields an unfolding curve in nanoDSF very similar to that of sLif in the presence of LipA but shifted compared to the unfolding curves of sLif$_{Y99A}$ and sLif. Dissociation constants obtained from fluorescence binding assay reveal approximately 3-fold stronger binding of sLif (K_D = 29 ± 9 nM) to pre-activated LipA than sLif$_{Y99A}$. The overall affinity is in the range previously reported for the complex from *B. glumae* (K_D = 5 nM) (10).

The combined analysis of time-resolved fluorescence anisotropy and polarization-resolved FCS yields parameters, rotational correlation times and order parameters, that can be used to investigate shape and local flexibility of the fluorescently labelled proteins.(20)

Here, we introduce a stochastic labeling strategy of lysine residues by the anisotropy sensor BDP FL, to obtain insights into protein conformational flexibility and motions on the single-molecule level. We showed that, within the resolution of our approach, the Y99A mutation does not affect the shape and mobility of free sLif (Fig. 3E, magenta and violet arrows). In light of these results, we expected that the mutation Y99A would not substantially alter the internal structure of sLif. This assumption was strengthened by our observation that the melting temperatures of the two proteins measured by nanoDSF were similar (48.1°C for sLif and 50.5°C for sLif$_{Y99A}$) (Fig. 3A). However, sLif only contains a single fluorescent residue (W283) located in MD2, such that one cannot exclude that the Y99A mutation in MD1 may lead to structural changes that are not detected in the distant MD2.

With this in mind, we aimed at solving the 3D structures of sLif and sLif$_{Y99A}$ by X-ray crystallography. Unfortunately, probably due to sLif's dynamic behavior,(7) all attempts have failed. We thus resorted to solving the structures of isolated *P. aeruginosa* MD1 and variant MD1$_{Y99A}$ using solution NMR spectroscopy (Fig. 4). The MD1 solution structure provides the first experimental evidence that this domain adopts a stable tertiary structure even in the absence of LipA. The structure folds as a three α-helical bundle, stabilized by hydrophobic and aromatic residues (I93, F97, F100, L116, I120, L135, M136 and Y139) and is very similar to the structure of MD1 from the *B. glumae* Lif:LipA complex (Fig. 4). Furthermore, structural comparison of MD1 and MD1$_{Y99A}$ revealed only minor structural changes, excluding that domain rearrangements or unfolding are the cause of the inactivation effect that is induced by the Y99A mutation. As such, the structural data alone does not allow to pinpoint the role of Y99 for LipA activation.

In the X-ray structure of *B. glumae* Lif:LipA complex,(10) residue Y91 (structurally equivalent to Y99 in *P. aeruginosa* Lif)

contributes only little (~110 Å² of solvent-accessible surface area) to the overall interface of ~5400 Å² formed between the two proteins, which involves 65 Lif residues. Hence, it is not unexpected that *P. aeruginosa* $s\text{Lif}_{Y99A}$ forms a complex with LipA, yet also this mutation decreases the affinity of *s*Lif towards LipA, which is confirmed by our fluorescence-based assay and NMR observation (Fig. S7). Indeed, we showed in *in vitro* experiments where isolated MD1 or MD1_{Y99A} compete with *s*Lif for LipA binding that MD1, but not MD1_{Y99A}, inhibits LipA activation (Fig. 2C). pFCS and fluorescence anisotropy decay also highlight the important difference in mechanics of *s*Lif and $s\text{Lif}_{Y99A}$ bound to LipA, where *s*Lif appears to be less mobile than in free form, while the local mobility of $s\text{Lif}_{Y99A}$ is increased compared to *s*Lif (Fig. 3E, orange and blue arrows). Altogether, this indicates that for the mutant Y99A the MD1 domain does not, or only weakly, interact with LipA while the rest of sLif still interacts normally. Since DSF cannot sense effects in the MD1 domain (no Trp), this interpretation is with all data. Importantly, the increased flexibility of complexed $s\text{Lif}_{Y99A}$ could also largely amplify the effect seen in the CNA.

Chemically denatured LipA, when refolded *in vitro*, adopts a globular pre-active conformation that shows a similar secondary structure content, intrinsic tryptophan fluorescence and susceptibility to proteolytic degradation as the native and active form of LipA.(7, 14) This pre-active intermediate is converted to active LipA by addition of Lif. Apparently, structures of pre-active and native LipA are very similar, however, it is still unknown which structural changes in LipA are caused by Lif during activation. By analyzing the structures of native LipA from *P. aeruginosa* and *B. glumae*(*9, 10, 23*) in different states, we observed conformational differences in a region close to the active site formed by residues 17-30 in lipases of *P. aeruginosa* and *B. glumae*. These residues are involved in the formation of a short, two-stranded β-sheet accompanying the opening of the lid-like α-helix covering the active site. We thus hypothesize that Lif may mediate LipA activation by promoting the formation of a short β-sheet formed by residues 17-30.

We tested this hypothesis by molecular simulations of *s*Lif and $s\text{Lif}_{Y99A}$. CNA perturbation analyses of conformational ensembles generated by MD simulations starting from the comparative model of *P. aeruginosa* Lif:LipA complexes revealed that substitution Y99A destabilizes residues 28-30 involved in the formation of the short, two-stranded β-sheet in LipA (Fig. 6). We also identified a long-range network of interactions involving residues 1-45, 197-202, 242-250 and 268-286 in LipA that span from Y99 of Lif to the LipA $\beta\text{-sheet}_{17\text{-}30}$ adjacent to the active site. We thus propose a mechanism of LipA activation based on the formation and stabilization of $\beta\text{-sheet}_{17\text{-}30}$ in LipA through interactions with Lif. Such small structural changes upon LipA activation are expected given the previously observed similarity of pre-active and native LipA.(7, 14)

In summary, our study reveals an intricate role of Y99 of the MD1 domain of *P. aeruginosa* Lif for LipA activation. Despite almost no influence on the global MD1 structure and weakly on *s*Lif global binding to LipA, the Y99A substitution hampers LipA activation, by the disruption of the mechanics of Lif:LipA complex. Molecular simulations suggest that, by long-range network interactions, Y99 supports the formation of a key secondary structure element in LipA on the way from pre-active to native LipA. Thus, our study for the first time provides insights at the atomistic level as to a potential mechanism of Lif-mediated pre-active-to-native LipA folding. This finding might spark further *in vitro* and *in vivo* studies to validate this putative mechanism.

Experimental procedures

Cloning, protein production and purification. The expression plasmid encoding MD1 (pET-MD1) (Tab. S6) of Lif was created by PCR using Phusion DNA polymerase (Thermo Fischer Scientific) in whole plasmid amplification with mutagenic oligonucleotides (Tab. S6) designed for SLIC method(27) and pEHTHis19(28) plasmid as a template. For that purpose, amino acids 1-65 in Lif were deleted using primers Lif_dLinkVD_fw and Lif_dLinkVD_rv leaving the amino acid sequence MGHHHHHH before amino acid L66 of Lif. Subsequently, the sequence behind amino acid L146 was removed with the same method using the primers Lif_backbone_fw and Lif_MD1_rv. The expression plasmids pET-MD1_{Y99A} and pET-$s\text{Lif}_{Y99A}$ respectively encoding MD1_{Y99A} and $s\text{Lif}_{Y99A}$, the variants with mutation of Y99 to alanine, were created by whole plasmid PCR amplification with mutagenic oligonucleotide pair Lif_Y99A_fw/Lif_Y99A_rv designed for SLIC

method and pET-MD1 or pEHTHis19 templates, respectively. *s*Lif, *s*Lif$_{Y99A}$, MD1 and MD1$_{Y99A}$ were expressed in *E. coli* BL21 (DE3) using respective expression plasmids as described previously.(28) For NMR studies, LB medium(29) was replaced by M9 media supplemented with ^{13}C-glucose and $^{15}NH_4Cl$ as sole carbon and nitrogen sources, respectively.

*s*Lif, *s*Lif$_{Y99A}$, MD1 and MD1$_{Y99A}$ variants carrying N-terminal His$_6$-tag were purified by immobilized metal affinity chromatography using Ni-NTA resins (Qiagen) according to a protocol of Hausmann *et al.* (2008).(28) Purified proteins were transferred to 20 mM sodium phosphate buffer (pH 7.4) by using PD10 column and concentrated to 500-1000 µM by using ultrafiltration device (Vivaspin) with a membrane of 5 kDa pore size. Protease inhibitor cocktail (0.1 x) (Sigma Aldrich) and 3 mM NaN_3 were added to the final samples in order to ensure protein stability during long-term NMR experiments.

Resonance assignment and structure calculation. NMR experiments were performed on Bruker Avance III HD$^+$ spectrometers operating either at 600 or 700 MHz, both equipped with 5 mm inverse detection triple-resonance z-gradient cryogenic probes. Data was collected at 30°C with sample concentrations between 450 and 900 µM in 20 mM sodium phosphate buffer pH 7.4 containing 10% (v/v) D_2O, 0.01% sodium azide and 100 µM 4,4-dimethyl-4-silapentanesulfonic acid (DSS). All NMR spectra were processed with TOPSPIN 3.5 (Bruker BioSpin). DSS was used as a chemical shift standard and ^{13}C and ^{15}N data were referenced using frequency ratios as previously described(30).

For the resonance assignment of MD1 and MD1$_{Y99A}$, ^{15}N- and ^{13}C-edited HSQC (heteronuclear single-quantum coherence) and three-dimensional HNCO, HN(CA)CO, HN(CO)CACB and HNCACB experiments were performed to obtain the chemical shift assignments of the backbone atoms. Three-dimensional ^{15}N- and ^{13}C-NOESY-HSQC and (H)CCH-TOCSY, spectra were used for side-chain resonance assignment and NOE (nuclear Overhauser effect) measurements using acquisition parameters listed in tables S2 and S3.

After assignment completion, CYANA2.1(*31*) was used to analyze the peak data from the NOESY spectra in a semi-automated iterative manner. We used CARA 1.9.24a(*32*) to automatically generate the NOE co-ordinates and intensities. The input data consisted of the amino acid sequence (to which we removed the histidine tag due to the lack of constraints), assigned chemical shift list, peak volume list and backbone dihedral angles (Φ and Ψ) derived from TALOS+ server(33) or with the CYANA script GridSearch.(31) The unambiguous NOEs assigned to a given pair of protons were converted into the upper limits by CYANA2.1.(31) No stereospecific assignments were introduced initially. In the final steps, 12 and 21 pairs of stereospecific restraints were introduced by CYANA2.1(31) for MD1 and MD1$_{Y99A}$, respectively.

The 20 conformers with the lowest final CYANA target function values were subjected to restrained energy-minimization as described in Pimenta *et al.* (2013)(34) with the AMBER14 software package using the ff14SB force field.(35) The structures were placed in an octahedral periodic box of TIP3P water molecules.(36) The restrained energy minimization was then performed in three stages. First, the solute atoms were kept fixed with harmonic positional restraints with a force constant of 500 kcal mol^{-1} Å^{-2} to relax the solvent molecules. Subsequently, the entire system was relaxed after restraint removal. During the last stage, 1500 steps of NMR-restrained energy minimization were applied with a combination of steepest descent minimization followed by conjugate gradient minimization. A parabolic penalty function was used for the NOE upper distance restraints with a force constant of 20 kcal mol^{-1} Å^{-2}. Finally, the geometric quality of the refined structures was analyzed with the Protein Structure Validation Software suite (version 1.5).(37) Statistics for the NMR solution structures of MD1 and MD1$_{Y99A}$ are given in table 1. The structural coordinates were deposited in the Protein Data Bank (PDB) under the accession codes 5OVM and 6GSF and the NMR data was deposited in the Biological Magnetic Resonance Data Bank (BMRB) under the accession numbers 34175 and 34286 for MD1 and MD1$_{Y99A}$, respectively.

MD1 and MD1$_{Y99A}$ backbone dynamics. To gain insight into the backbone dynamics of MD1 and MD1$_{Y99A}$ in the solution we measured the relaxation parameters R_1, R_2 and {^{1}H}-^{15}N-NOE (HetNOE) for both proteins at 35°C. We used ^{15}N-labelled samples at a concentration of 650 and 600 µM for MD1 and MD1$_{Y99A}$, respectively. The solutions were prepared either in Tris-Glycine buffer pH 9 containing 10%

(v/v) D_2O, 0.01% sodium azide and 100 µM DSS or in 20 mM sodium phosphate buffer pH 7.4 containing 10% (v/v) D_2O, 0.01% sodium azide and 100 µM DSS for MD1 and $MD1_{Y99A}$, respectively. All data were collected in a Bruker Avance III HD^+ 600 MHz spectrometer.

Backbone relaxation rates, R_1 and R_2, were determined by acquiring pseudo-3D spectra consisting of a series of 2D heteronuclear ^1H-^{15}N-HSQC experiments were the relaxation period varied. For the ^{15}N longitudinal relaxation rates (R_1), 12 time points were collected (0.02s, 0.06s, 0.1s, 0.2s, 0.4s, 0.5s, 0.6s, 0.7s, 0.8s, 1.2s, 1.5s and 2s). The spectra were acquired with 2048 points in ^1H indirect dimension and 128 points in the ^{15}N direct dimension and 8 scans. The spectral width was 7183.9 Hz in the ^1H dimension and 1943.8Hz in the ^{15}N dimension and the relaxation delay was 1.5s. The central frequency for proton was set on the solvent signal (2812.9 Hz) and for nitrogen was set on the center of the amide region (7535.96 Hz). For the ^{15}N transverse relaxation rate (R_2) 12 time points were collected (0.02s, 0.03s, 0.04s, 0.06s, 0.08s, 0.1s, 0.12s, 0.14s, 0.16s, 0.2s, 0.24s and 0.28s). The spectra were acquired in the same conditions as the above. The $\{^1H\}$-^{15}N-NOE steady-state NOE experiments were recorded with a relaxation delay of 10 s, with 8 transients in a matrix with 2048 data points in F2 and 256 increments in F1, in an interleaved manner, with alternating proton-pre-saturated and non-pre-saturated spectra. All data was processed with TopSpin 3.5 (Bruker BioSpin) and analyzed with CARA 1.9.24a(32) and Relax 4.0.3.(38)

MD1 and $MD1_{Y99A}$ interaction with LipA. To investigate the effects of the mutation on the interaction with LipA, we followed the backbone signal intensity in ^1H,^{15}N-HSQC spectra of 120 µM ^{15}N-labeled MD1 or $MD1_{Y99A}$ in the presence and absence of 400 µM (unlabeled) LipA. All data was acquired at 10°C, processed and analyzed with TopSpin 3.5 (Bruker BioSpin). Samples were prepared in Tris-Glycine buffer pH 9 containing 10% (v/v) D_2O, 0.01% sodium azide and 100 µM DSS.

In vitro activation of LipA with Lif. LipA, comprising residues S26-L311 without any affinity tag, was expressed in *E. coli* (BL21) DE3 using the plasmid pLipA-SS.(28) Cells expressing insoluble inclusion bodies of LipA were suspended in Tris-HCl buffer (100 mM, pH 7) containing 5 mM EDTA and 1 mM TCEP and disrupted by a French press. LipA inclusion bodies were collected by centrifugation at 10,000 *g* for 10 min and suspended in the same buffer. Centrifugation and wash steps were repeated three times to obtain purified LipA inclusion bodies. These were solubilized with Tris-HCl buffer containing 8 M urea at 37°C for 1 h and remaining insoluble material was removed by centrifugation at 10,000 *g* for 10 min. Solubilized LipA was refolded by dilution with the TG buffer (5 mM glycine, 5 mM Tris, pH 9) containing an equimolar amount of *s*Lif and was incubated overnight at 4°C.

Lipase activity assay. The activity of LipA was spectrophotometrically monitored by the release of *p*-nitrophenolate from the standard lipase substrate *p*-nitrophenyl palmitate (*p*NPP, 1 mM) in 10 mM TG buffer containing 1 mM $CaCl_2$.(1)

Inhibition of LipA activation. LipA (50 nM) was incubated with MD1 or $MD1_{Y99A}$ (0.2 – 1.0 µM) for 1 h at room temperature in a glass-coated 96-well microtiter plate (MTP) followed by addition of *s*Lif (50 nM). After agitating for 10 min at room temperature, *p*NPP lipase substrate (100 µL) was added to 100 µL of activated LipA in MTP and lipase activity was determined.

Co-purification assays. In vitro Lif-LipA interaction studies were performed using *s*Lif and $sLif_{Y99A}$ variants carrying N-terminal His_6-tags for immobilization onto Ni-NTA resins. First, the complexes of LipA (4 µM) with *s*Lif (4 µM) or $sLif_{Y99A}$ (4 µM) were formed in Tris-HCl buffer (10 mM Tris-HCl, pH 9) by incubation overnight at 4°C followed by loading onto a Ni-NTA column and exhaustively washing with Tris-HCl buffer (10 mM, pH 9). The proteins bound to the column were eluted with Tris buffer (10 mM, pH 9) containing 500 mM imidazole. Elution fractions were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) under denaturation conditions on 16 % (w/v) gels(39) followed by staining with Coomassie Brilliant Blue G250.

Protein stability determination by differential scanning fluorimetry. LipA (2 µM) was incubated with *s*Lif (2 µM), $sLif_{Y99A}$ (2 µM), MD1 (2 µM) or $MD1_{Y99A}$ (2 µM) overnight at 4°C in TG buffer. The protein samples loaded into the measuring capillaries (Prometheus NT.Plex nanoDSF Grade Standard Capillary Chips) were heated from 15°C to 95°C (heating rate of 0.2°C/min) and the intrinsic protein fluorescence was recorded at 330 nm and 350 nm using the Prometheus NT.Plex nanoDSF

device (Nano Temper, Munich, Germany). The ratio of F_{350nm} and F_{330nm} and its first derivative were calculated by the PR.ThermControl software provided by the company.

Model building of the P. aeruginosa Lif:LipA complex. The three-dimensional structure of *P. aeruginosa* Lif is currently unknown. Thus, a homology model of the *P. aeruginosa* *s*Lif:LipA complex was constructed using the structure of the *B. glumae* lipase:foldase complex (PDB code 2ES4) as a template (sequence identity/similarity: 39%/52%) for Lif and 41%/73% for LipA). The Phyre2 web server(40) was used for homology modelling. The model obtained was energy minimized with the GROMOS96 43B1 force field as implemented in Swiss-PdbViewer.(41) After ten rounds of energy minimization, the C_α atoms of the models were superimposed on the template structure and the model with the lowest RMSD was taken for further studies. The model obtained was evaluated by using our in-house model quality assessment program TopScore.(24, 25) The correctness of the model is measured as the predicted global and local lDDT score(42) compared to the native structure. The lDDT score compares all intra-molecular heavy-atom distances within two structures and, thus, is superposition-free. Two models are considered entirely different if all distances deviate by more than 4 Å and completely identical if all distances deviate by less than 0.5 Å. Since the native structure is unknown, the score is predicted by a deep neural network which uses multiple sources of information as input. These include knowledge-based angle, distance and contact potentials, assessment of residue stereochemistry and atom clashes, model clustering and agreement between features predicted from the sequence and measured in the model, such as secondary structure, solvent accessibility and residue contacts. The deep neural network was trained on a large data-set of 660 protein targets totaling over 1.33×10^5 models and over 1.9×10^7 residues. The *P. aeruginosa* and *B. glumae* lipase:foldase complex structures show structural conservation of functionally important residues (Fig. S1), as for example the foldase motif residues (RXXFDY(F/C)L(S/T)A, X represents any amino acid) important for lipase activation(19) and R343 related to the specificity of *B. glumae* foldase to bind its cognate lipase.(10) To validate our complex model, we mapped all conserved amino acids to the structures and found all of them at the interface of LipA and Lif, as expected.

Molecular dynamics simulations. The refined model of *P. aeruginosa* *s*Lif:LipA complex was used as input structure for MD simulations. All-atom MD simulations were performed with the Amber 11 software package(*43*) using the ff99SB force field(44) as described in Ciglia *et al.* (2015).(45) The *s*Lif:LipA complex was placed in an octahedral periodic box of TIP3P water molecules(36) such that the smallest distance between the edges of the box and the closest solute atom is 11 Å. The SHAKE algorithm(45) was applied to constrain bond lengths of hydrogen atoms and long-range electrostatic interactions were taken into account using the Particle Mesh Ewald method.(46) The time step was set to 2 fs with a non-bonded cut-off of 8 Å. The starting structures were first energy minimized by applying 50 steps of steepest descent minimization, followed by 450 steps of conjugate gradient minimization. During the minimization, the solute atoms were restrained applying decreasing harmonic potentials, with a force constant of 25 kcal mol^{-1} $Å^{-2}$ initially, reduced to 5 kcal mol^{-1} $Å^{-2}$ in the last round. For thermalization, the systems were heated from 100 K to 300 K in 50 ps of canonical (NVT)–MD simulations applying harmonic potentials with a force constant of 5 kcal mol^{-1} $Å^{-2}$ on the solute atoms. Afterwards, MD simulations of 250 ps length were performed in the isothermal-isobaric ensemble (NPT) with the same harmonic potentials to adjust the density of the simulation box. Finally, the force constant of the harmonic restraints was reduced to zero during 100 ps of MD simulations in the NVT ensemble. For production, ten independent, unbiased MD simulations of 1.5 µs length were performed, totaling 15 µs of production runs. To ensure independence, production runs were carried out at temperatures of 300.0 K + T, where T was varied from 0.0 to 0.9 for each run, respectively.

*Constraint network analysis (*CNA*).* To detect changes in *s*Lif:LipA rigidity and flexibility upon Y99A mutation in *P. aeruginosa* Lif, we analyzed an ensemble of snapshots of *s*Lif-bound LipA in terms of a perturbation approach(26) in a similar way as done by Milić *et al.*(25) In short, first, an ensemble of 7,500 constraint network topologies was generated from MD snapshots of the proteins sampled at 2 ns intervals from the 10 MD simulations of the *s*Lif:LipA complex (see above). Second, altered

bimolecular stability due to the *s*Lif$_{Y99A}$ mutation is quantified in terms of a per-residue decomposition of the perturbation free energy $\Delta G_{i,\mathrm{CNA}}$, following a linear response approximation (eq. 1).

$$\Delta G_{i,\mathrm{CNA}} = \alpha\left(\langle E_{i,\mathrm{CNA}}^{perturbed}\rangle - \langle E_{i,\mathrm{CNA}}^{ground}\rangle\right) \qquad \textbf{eq. 1}$$

Parameter α has been generally determined empirically and was set to 0.02 as in Pfleger *et al.*(*26*) $\Delta G_{i,\mathrm{CNA}}$ was computed based on rigidity analyses performed with the CNA software package(*26*) on ensembles of network topologies of the ground (*s*Lif) and perturbed (*s*Lif$_{Y99A}$) states. Upon perturbation, about 19% of the residues in *s*Lif and 22% of the residues in LipA show altered stability characteristic according to $\Delta G_{i,\mathrm{CNA}}$ values > 0.1 kcal mol^{-1}.

Fluorescence labelling of Lif. Purified proteins *s*Lif, *s*Lif$_{Y99A}$, MD1 and MD1$_{Y99A}$ were transferred to 50 mM sodium phosphate buffer (pH 7.4) and concentration was adjusted to 70 µM. To label amino groups, Bodipy FL NHS ester (BDP FL; Lumiprobe), was dissolved in DMSO and added to the protein in 1:10 molar ratio to ensure labeling of single dye per protein molecule (obtained degree of labeling approximately 5%). Free dye was removed after overnight incubation at 4°C by buffer exchange with Amicon Ultra-0.5 mL 10K centrifugal filters (Merck-Milipore).

Fluorescence measurements and data analysis. Steady-state fluorescence anisotropy $r_{\text{steady-state}}$ and average translational diffusion time $\langle t_{trans}\rangle$ were measured in the droplets on a cover-slide for 20 seconds to avoid changes of LipA concentrations due to protein adsorption. The concentrations of the labeled proteins were in the range of 1.2 ± 0.1 nM. The fluorescence signal was recorded on a custom-built confocal microscope(47) with polarization-resolved detection with parallel- and perpendicular-polarized channels, $F_p(t)$ and $F_s(t)$. Anisotropy was calculated using equation:

$$r_{steady-state} = F_p - G_f \cdot F_s / F_p + G_f \cdot 2F_s \qquad \textbf{eq. 2}$$

where G_f is the detection efficiency ratio between parallel and perpendicular channel.

The average translational diffusion time t_{dif} was calculated using Software Package for Multi-parameter Fluorescence Spectroscopy, Full Correlation and Multi-parameter Fluorescence Imaging.(48) Correlation curves $G(t_c)$ were approximated with 3-dimensional Gaussian diffusion model with 2 photophysical bunching terms:

$$G(t_c) = \frac{1}{N}\left(1+\frac{t_c}{\langle t_{trans}\rangle}\right)^{-1}\left(1+\left(\frac{\omega_0}{z_0}\right)^2 \times \frac{t_c}{\langle t_{trans}\rangle}\right)^{-\frac{1}{2}} + \left(1 - b_1 + b_1 e^{-t_c/t_{b1}} - b_2 + b_2 e^{-t_c/t_{b2}}\right) \qquad \textbf{eq. 3}$$

Here, the observation volume is approximated by a 3D-Gaussian volume with $1/e^2$ radii in the lateral (ω_0) and axial direction (z_0), with the particle number N, $\langle t_{trans}\rangle$ is the apparent average translational diffusion time for the free and complexed sLif and MD1 variants, respectively, $b_{1,2}$ and $t_{b1,b2}$ are amplitudes and times of the bunching terms.

The fraction of the complex $x_{complex}$ was obtained from the linear combination of the fluorescence parameters of free *s*Lif-BDP FL and of *s*Lif-BDP FL in presence of >10 µM of LipA, assigned to be associated with Lif-LipA complex.

Polarization-resolved full fluorescence correlation spectroscopy was performed with a confocal laser scanning microscope (FV1000, Olympus, Germany) equipped with a single photon counting device with picosecond time-resolution (4 detectors, PD5CTC, Micro Photon Devices, Bolzano, Italy; counting electronics, HydraHarp400, PicoQuant, Berlin, Germany) at 23.5 ± 0.5°C. The sample was excited by the continuous wave parked beam at 488 nm and the fluorescence F was collected in s- and p-polarized channels, $F_s(t)$ and $F_p(t)$, respectively. Full cross-correlation curves, $G_{p,p}(t_c)$ and $G_{s,s}(t_c)$, $G_{s,p}(t_c)$ and $G_{p,s}(t_c)$, were obtained according to Felekyan *et al.*(48) Data were processed as previously described in Möckel *et al.*(20)

Time-resolved fluorescence anisotropy decay curves were recorded using a FluoTime300 fluorescence lifetime spectrometer (PicoQuant, Berlin, Germany) equipped with a pulsed super continuum laser SuperK Extreme (NKT Photonics, Denmark) as a light source running at 15.61 MHz and a wavelength of 488 nm in a temperature-stabilized cell at 20.0 ± 0.1°C. The fluorescence and anisotropy decays were recovered by global fitting of the sum ($F_p(t_c)$+$F_s(t_c)$) and difference ($F_p(t_c)$-$F_s(t_c)$) curves as previously described (eq. S2a, b).(20)

In vitro binding of LipA and fluorescently labelled sLif.

Labelled *s*Lif/ *s*Lif$_{Y99A}$ –BDP FL (1.2 ± 0.1 nM of BFL, total concentration c^0_{sLif} approximately 24 nM)) was incubated overnight at 4°C with various concentrations of LipA (0 – 50 µM) in 10 mM glycine buffer (10 mM, pH 9) in protein low binding tubes (Sarstedt AG). Equilibrium dissociation constant K_D was fitted using 1:1 binding model:

$$x_{complex} = \frac{1}{c^0_{sLif}}\left(\frac{K_D + c^0_{sLif} + c^0_{LipA}}{2} - \sqrt{\left(\frac{K_D + c^0_{sLif} + c^0_{LipA}}{2}\right)^2 - c^0_{sLif} \cdot c^0_{LipA}}\right) \quad \textbf{eq. 4}$$

where $x_{complex}$ is a fraction of Lif in complex with LipA, KD is a dissociation constant and $c^0_{sLif}$and $c^0_{LipA}$are the total concentrations of *s*Lif and LipA.

Acknowledgements: AV and ME acknowledge access to the Jülich-Düsseldorf Biomolecular NMR Center. TV acknowledges support from the International North-Rhine-Westphalia Research School iGRASPseed. H.G. is grateful for computational support and infrastructure provided by the "Zentrum für Informations- und Medientechnologie" (ZIM) at the Heinrich Heine University Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) to HG on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user ID: HDD16, project IDs: 11237, 12732).

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Jaeger, K. E., and Kovacic, F. (2014) Determination of lipolytic enzyme activities, *Methods Mol Biol 1149*, 111-134.
2. Bleves, S., Viarre, V., Salacha, R., Michel, G. P., Filloux, A., and Voulhoux, R. (2010) Protein secretion systems in *Pseudomonas aeruginosa*: A wealth of pathogenic weapons, *Int J Med Microbiol 300*, 534-543.
3. Jaeger, K. E., Kharazmi, A., and Hoiby, N. (1991) Extracellular lipase of *Pseudomonas aeruginosa*: biochemical characterization and effect on human neutrophil and monocyte function in vitro, *Microb Pathog 10*, 173-182.
4. Konig, B., Jaeger, K. E., and Konig, W. (1994) Induction of inflammatory mediator release (12-hydroxyeicosatetraenoic acid) from human platelets by *Pseudomonas aeruginosa*, *Int Arch Allergy Immunol 104*, 33-41.
5. Konig, B., Jaeger, K. E., Sage, A. E., Vasil, M. L., and Konig, W. (1996) Role of *Pseudomonas aeruginosa* lipase in inflammatory mediator release from human inflammatory effector cells (platelets, granulocytes, and monocytes, *Infect Immun 64*, 3252-3258.
6. Jaeger, K. E., Ransac, S., Dijkstra, B. W., Colson, C., van Heuvel, M., and Misset, O. (1994) Bacterial lipases, *FEMS Microbiol Rev 15*, 29-63.
7. Pauwels, K., Sanchez del Pino, M. M., Feller, G., and Van Gelder, P. (2012) Decoding the folding of *Burkholderia glumae* lipase: folding intermediates en route to kinetic stability, *PLoS One 7*, e36999.
8. El Khattabi, M., Ockhuijsen, C., Bitter, W., Jaeger, K. E., and Tommassen, J. (1999) Specificity of the lipase-specific foldases of Gram-negative bacteria and the role of the membrane anchor, *Mol Gen Genet 261*, 770-776.
9. Noble, M. E., Cleasby, A., Johnson, L. N., Egmond, M. R., and Frenken, L. G. (1993) The crystal structure of triacylglycerol lipase from *Pseudomonas glumae* reveals a partially redundant catalytic aspartate, *FEBS Lett 331*, 123-128.
10. Pauwels, K., Lustig, A., Wyns, L., Tommassen, J., Savvides, S. N., and Van Gelder, P. (2006) Structure of a membrane-based steric chaperone in complex with its lipase substrate, *Nat Struct Mol Biol 13*, 374-375.
11. Shibata, H., Kato, H., and Oda, J. (1998) Molecular properties and activity of amino-terminal truncated forms of lipase activator protein, *Biosci Biotechnol Biochem 62*, 354-357.
12. Rosenau, F., Tommassen, J., and Jaeger, K. E. (2004) Lipase-specific foldases, *Chembiochem 5*, 152-161.
13. Nagradova, N. (2007) Enzymes catalyzing protein folding and their cellular functions, *Curr Protein Pept Sci 8*, 273-282.

14. El Khattabi, M., Van Gelder, P., Bitter, W., and Tommassen, J. (2000) Role of the lipase-specific foldase of *Burkholderia glumae* as a steric chaperone, *J Biol Chem 275*, 26885-26891.
15. Ihara, F., Okamoto, I., Akao, K., Nihira, T., and Yamada, Y. (1995) Lipase modulator protein (LimL) of *Pseudomonas* sp. strain 109, *J Bacteriol 177*, 1254-1258.
16. Hobson, A. H., Buckley, C. M., Jorgensen, S. T., Diderichsen, B., and McConnell, D. J. (1995) Interaction of the *Pseudomonas cepacia* DSM3959 lipase with its chaperone, LimA, *J Biochem 118*, 575-581.
17. Eder, J., Rheinnecker, M., and Fersht, A. R. (1993) Folding of subtilisin BPN': characterization of a folding intermediate, *Biochemistry 32*, 18-26.
18. Dodson, K. W., Jacob-Dubuisson, F., Striker, R. T., and Hultgren, S. J. (1993) Outer-membrane PapC molecular usher discriminately recognizes periplasmic chaperone-pilus subunit complexes, *Proc Natl Acad Sci U S A 90*, 3670-3674.
19. Shibata, H., Kato, H., and Oda, J. (1998) Random mutagenesis on the *Pseudomonas* lipase activator protein, LipB: exploring amino acid residues required for its function, *Protein Eng 11*, 467-472.
20. Mockel, C., Kubiak, J., Schillinger, O., Kuhnemuth, R., Della Corte, D., Schroder, G. F., Willbold, D., Strodel, B., Seidel, C. A. M., and Neudecker, P. (2019) Integrated NMR, fluorescence, and molecular dynamics benchmark study of protein mechanics and hydrodynamics, *J Phys Chem B 123*, 1453-1480.
21. Tsytlonok, M., Sanabria, H., Wang, Y., Felekyan, S., Hemmen, K., Phillips, A. H., Yun, M. K., Waddell, M. B., Park, C. G., Vaithiyalingam, S., Iconaru, L., White, S. W., Tompa, P., Seidel, C. A. M., and Kriwacki, R. (2019) Dynamic anticipation by Cdk2/Cyclin A-bound p27 mediates signal integration in cell cycle regulation, *Nat Commun 10*, 1676.
22. Ortega, A., Amoros, D., and Garcia de la Torre, J. (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models, *Biophys J 101*, 892-898.
23. Nardini, M., Lang, D. A., Liebeton, K., Jaeger, K. E., and Dijkstra, B. W. (2000) Crystal structure of *Pseudomonas aeruginosa* lipase in the open conformation. The prototype for family I.1 of bacterial lipases, *J Biol Chem 275*, 31219-31225.
24. Mulnaes, D., and Gohlke, H. (2018) TopScore: Using deep neural networks and large diverse data sets for accurate protein model quality assessment, *J Chem Theory Comput 14*, 6117-6126.
25. Milic, D., Dick, M., Mulnaes, D., Pfleger, C., Kinnen, A., Gohlke, H., and Groth, G. (2018) Recognition motif and mechanism of ripening inhibitory peptides in plant hormone receptor ETR1, *Sci Rep 8*, 3890.
26. Pfleger, C., Minges, A., Boehm, M., McClendon, C. L., Torella, R., and Gohlke, H. (2017) Ensemble- and rigidity theory-based perturbation approach to analyze dynamic allostery, *J Chem Theory Comput 13*, 6343-6357.
27. Li, M. Z., and Elledge, S. J. (2007) Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC, *Nat Methods 4*, 251-256.
28. Hausmann, S., Wilhelm, S., Jaeger, K. E., and Rosenau, F. (2008) Mutations towards enantioselectivity adversely affect secretion of *Pseudomonas aeruginosa* lipase, *FEMS Microbiol Lett 282*, 65-72.
29. Bertani, G. (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*, *J Bacteriol 62*, 293-300.
30. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects, *J Biomol NMR 5*, 67-81.

31. Guntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods Mol Biol 278*, 353-378.
32. Rochus, L. J. K. (2004) *The Computer Aided Resonance Assignment Tutorial*, CANTINA Verlag, Goldau, Switzerland.
33. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J Biomol NMR 13*, 289-302.
34. Pimenta, J., Viegas, A., Sardinha, J., Martins, I. C., Cabrita, E. J., Fontes, C. M., Prates, J. A., and Pereira, R. M. (2013) NMR solution structure and SRP54M predicted interaction of the N-terminal sequence (1-30) of the ovine Doppel protein, *Peptides 49*, 32-40.
35. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB, *J Chem Theory Comput 11*, 3696-3713.
36. Jorgensen, W. L., Chandrasekhar, J., and Madura, J. D. (1983) Comparison of simple potential functions for simulating liquid water, *The Journal of Chemical Physics 79*, 926-935.
37. Bhattacharya, A., Tejero, R., and Montelione, G. T. (2007) Evaluating protein structures determined by structural genomics consortia, *Proteins 66*, 778-795.
38. Bieri, M., d'Auvergne, E. J., and Gooley, P. R. (2011) relaxGUI: a new software for fast and simple NMR relaxation data analysis and calculation of ps-ns and mus motion of proteins, *J Biomol NMR 50*, 147-155.
39. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature 227*, 680-685.
40. Kelley, L. A., and Sternberg, M. J. (2009) Protein structure prediction on the Web: a case study using the Phyre server, *Nat Protoc 4*, 363-371.
41. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, *Electrophoresis 18*, 2714-2723.
42. Mariani, V., Biasini, M., Barbato, A., and Schwede, T. (2013) lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests, *Bioinformatics 29*, 2722–2728.
43. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs, *J Comput Chem 26*, 1668-1688.
44. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters, *Proteins 65*, 712-725.
45. Ciglia, E., Vergin, J., Reimann, S., Smits, S. H. J., Schmitt, L., Groth, G., and Gohlke, H. (2014) Resolving hot spots in the C-terminal dimerization domain that determine the stability of the molecular chaperone Hsp90, *PLoS One 9*.
46. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald - an N.Log(N) method for Ewald sums in large systems, *J Chem Phys 98*, 10089-10092.
47. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J., and Seidel, C. A. M. (2010) Accurate single-molecule FRET studies using multiparameter fluorescence detection, *Method Enzymol 475*, 455-514.
48. Felekyan, S., Kuhnemuth, R., Kudryavtsev, V., Sandhagen, C., Becker, W., and Seidel, C. A. M. (2005) Full correlation from picoseconds to seconds by time-resolved and time-correlated single photon detection, *Rev Sci Instrum 76*.

49. Lang, D., Hofmann, B., Haalck, L., Hecht, H. J., Spener, F., Schmid, R. D., and Schomburg, D. (1996) Crystal structure of a bacterial lipase from *Chromobacterium viscosum* ATCC 6918 refined at 1.6 angstroms resolution, *J Mol Biol 259*, 704-717.

FOOTNOTES

This study was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through funding no. ET 103/2-1 to ME; no. SE 1195/16-1 to CAMS, no. GO 1367/1-1 to HG and no. JA 448/8-1 to KEJ. Part of the work of FK, KEJ and CAMS was funded by CRC 1208 (projects A02 and A08). AV received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement N° 660258.

The abbreviations used are: BDP FL, Bodipy FL NHS ester, 4,4-difluoro-4-bora-3a,4a-diaza-s-indacene; BMRB, Biological Magnetic Resonance Data Bank; CFD, catalytic folding domain; CNS, constraint network analysis; DSS, 4,4-dimethyl-4-silapentanesulfonic acid; EHD, extended helical domain; HSQC, heteronuclear single-quantum coherence; MD, molecular dynamics; MD1, mini-domain 1; MD2, mini-domain 2; MTP, microtiter plate; NOE, nuclear Overhauser effect; NPT, isothermal-isobaric ensemble; PDB, Protein Data Bank; *p*FCS, polarization resolved fluorescence correlation spectroscopy; *p*NPP, *p*-nitrophenyl palmitate; $r(t_c)$, time-resolved fluorescence anisotropy; $r_{\text{steady-state}}$, steady-state fluorescence anisotropy; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TMD, transmembrane domain; VLD, variable linker domain.

Table 1: Statistics for the NMR structures of MD1 and $MD1_{Y99A}$.

NMR distance and dihedral constraints	MD1	$MD1_{Y99A}$
Distance constraints		
Total distance restraints from NOEs	964	1062
Short-range (|i−j|=1)	676	694
Medium-range (|i−j|<5)	199	264
Long-range (|i−j|≥5)	89	104
Total dihedral restraints	104	219
Structure statistics		
Violations (mean)		
Distance constraints (Å)	0.0042±0.0015	0.0055±0.0015
Dihedral angle constraints (Å)	0.3001±0.1139	0.2688±0.0779
CYANA target functions, Å	0.70±0.08	1.02±0.15
Average pairwise rmsd for residues 10-85		
Backbone atoms	1.45±0.22	1.37±0.26
Heavy atoms	1.84±0.16	1.88±0.22
Ramachandran's plot analysis		
Most favored regions	93.0%	90.4%
Additionally allowed regions	7.0%	6.7%
Generously allowed regions	0.0%	2.3%
Disallowed regions	0.0%	0.0%

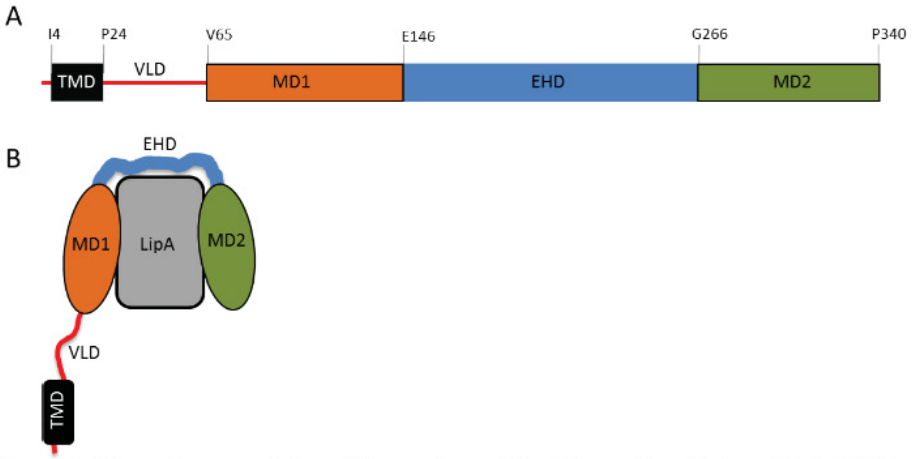


Figure 1. Schematic representation of *P. aeruginosa* Lif and its complex with lipase LipA. (A) Five-domain organization of Lif and (B) Lif-LipA complex. The catalytic folding domain (CFD) self-sufficient for activation of LipA *in vitro* comprises MD1, EHD and MD2. Residues defining the beginning and the end of each domain are indicated in (A).

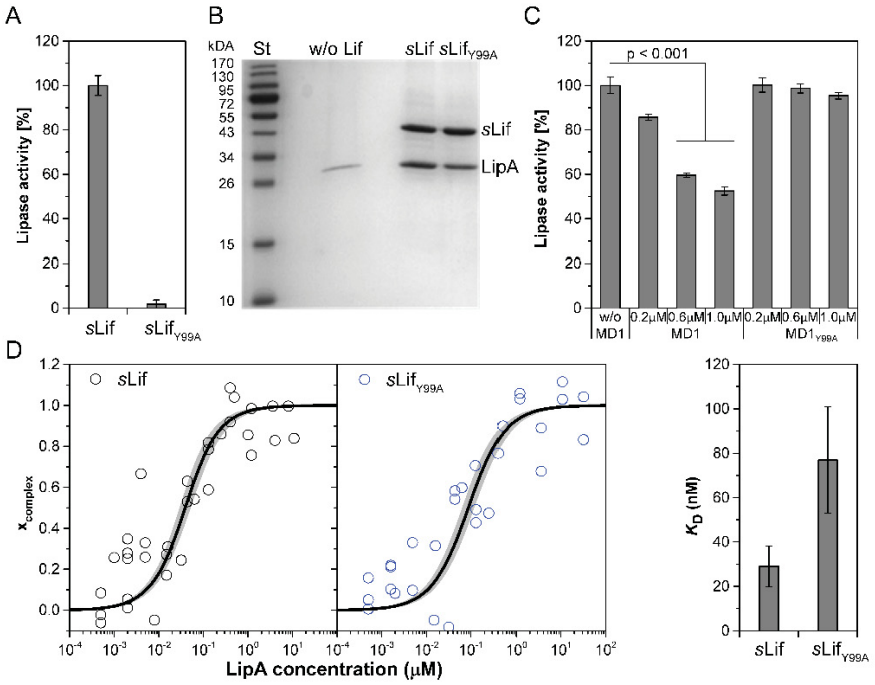


Figure 2. Activation of *P. aeruginosa* lipase LipA with *s*Lif and variant *s*Lif$_{Y99A}$ and effect of MD1. (A) Pre-active LipA (4 μM) was incubated with either *s*Lif (4 μM) or variant *s*Lif$_{Y99A}$ (4 μM) followed by lipase activity assay with 10 nM LipA. The activity of *s*Lif:LipA complex was set as 100%. (B) SDS-PAGE analysis of LipA co-purified in the complex with *s*Lif or *s*Lif$_{Y99A}$ as well as without Lif (w/o Lif). In the Coomassie Brilliant Blue G250 stained gel LipA is migrating as ~30 kDa and *s*Lif as 43 kDa protein. Molecular weights of standard proteins (St) are indicated on the left-hand side. (C) Inhibition of *s*Lif-mediated LipA activation with MD1. Pre-active LipA (50 nM) incubated with MD1 or MD1$_{Y99A}$ was activated by addition of *s*Lif (50 nM) and 10 min incubation prior to lipase activity measurement. Lipase activities are mean values ± standard deviation of three independent experiments each measured with at least three samples. (D) A fluorescence assay was used to study the complex formation of pre-active LipA and *s*Lif/*s*Lif$_{Y99A}$ labelled at amino groups with BDP FL. The fraction of fluorescence parameters assigned to the *s*Lif:LipA complex (steady-state anisotropy $r_{steady\text{-}state}$ (eq. 2, Tab. S1) and average translational diffusion time $\langle t_{trans} \rangle$ (eq. 3, Tab. S2).The binding data were fitted with a 1:1 binding affinity model (eq, 4), black line. The uncertainties are indicated as shaded areas. Steady-state anisotropy could not be used for *s*Lif$_{Y99A}$:LipA complex because increased mobility of the fluorescent probe cancels the increase of global rotation correlation time ρ_{global}. The apparent dissociation constant K_D (right panel) was determined to 29 nM ± 9 nM for *s*Lif:LipA and 77 nM ± 24 nM for *s*Lif$_{Y99A}$:LipA complex (error bars are standard errors of the fit).

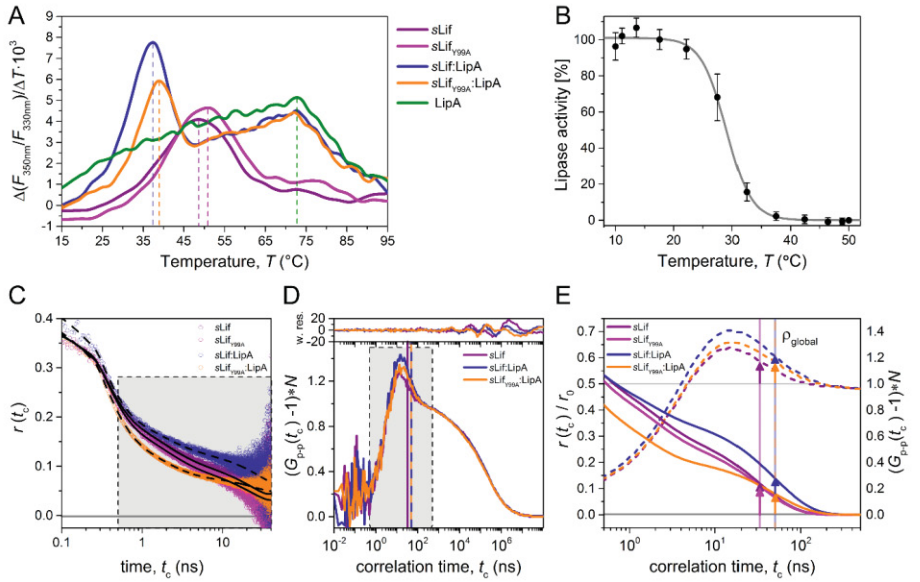


Figure 3. Effect of temperature on unfolding and activity of LipA in complex with *s*Lif and variant *s*Lif$_{Y99A}$. (A) Melting curves of *s*Lif, *s*Lif$_{Y99A}$, pre-active LipA alone and after incubation with *s*Lif, *s*Lif$_{Y99A}$, MD1 and MD1$_{Y99A}$ obtained by fluorescence measurement with nanoDSF. Green dashed lines indicate melting temperatures of LipA in complex with *s*Lif and *s*Lif$_{Y99A}$, blue dashed lines indicate melting temperatures of *s*Lif and *s*Lif$_{Y99A}$ and the melting temperature of pre-active LipA is indicated with the orange dashed line. (B) Temperature-dependent lipase activity in a solution of sLif and LipA generated by incubation of pre-active LipA (100 nM) with *s*Lif (250 nM) overnight at 4°C in TG buffer. Samples were then incubated at different temperatures (10 – 50°C) for 1 h followed by measurement of the remaining lipase activity with 2 nM LipA. (C) Time-resolved fluorescence anisotropy decay curves $r(t_c)$ of fre and complex sLif and *s*Lif$_{Y99A}$. Open circles indicate experimental $r(t_c)$ and lines indicate model $r(t_c)$ (eq. S2b, results see Tab. S4), dashed lines indicate complex. Further details see main text. (D) Polarization-resolved full-FCS of labelled *s*Lif using the p-p cross-correlation curves $G_{p\text{-}p}(t_c)$ normalised to the number of molecules in focal volume (eq. S3) together with weighted residuals of the fits (upper plot, eq. S4, results in Tab. S5). The global rotation correlation time ρ_{global} of *s*Lif is similar to the one obtained by anisotropy measurement (33 ns, indicated by vertical line). The global rotation correlation times of *s*Lif:LipA and *s*Lif$_{Y99A}$:LipA are similar (50 ns, vertical line). (E) Joint analysis of the anisotropy order parameters (solid lines, see shaded area in C) and normalized pFCS amplitudes (dashed lines, see shaded area in D) by displaying the model functions of the fits. The global rotational correlation times are depicted as vertical lines. The corresponding amplitudes are highlighted by arrows. Further details see main text.

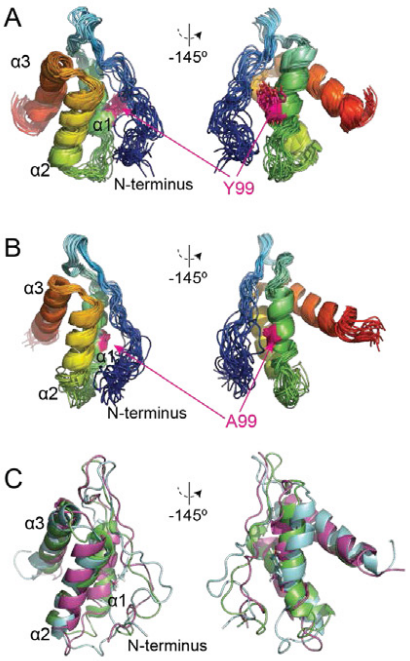


Figure 4. Details of MD1 and variant $MD1_{Y99A}$ structures obtained by NMR spectroscopy. (A) Cartoon representations of the structure ensemble of the 20 best solution structures of MD1 and (B) $MD1_{Y99A}$ variant. (C) Comparison of the representative NMR solution structures of MD1 (cyan) and $MD1_{Y99A}$ (purple) with the crystal structure of MD1 from *B. glumae* (green) (PDB code 2ES4(10)).

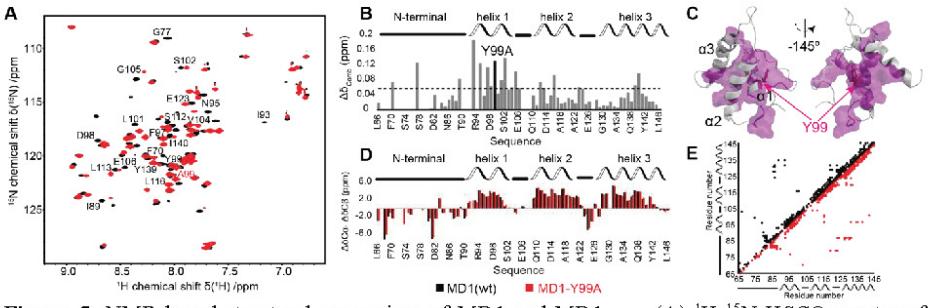


Figure 5. NMR-based structural comparison of MD1 and $MD1_{Y99A}$. (A) ^{1}H-^{15}N-HSCQ spectra of MD1 (black) and $MD1_{Y99A}$ (red). Labels correspond to the most affected residues due to the mutation. (B) Chemical shift perturbations induced by the mutation along the MD1 sequence and (C) mapped on the MD1 structure (purple, mutation site highlighted). (D) Comparison of ^{13}C secondary chemical shifts of MD1 (black) and $MD1_{Y99A}$ (red). Positive/negative values indicate α-helical/β-strand secondary structure. Random coil values should be zero. (E) Inter-residue distance restraints from NOEs for MD1 (black) and $MD1_{Y99A}$ (red).

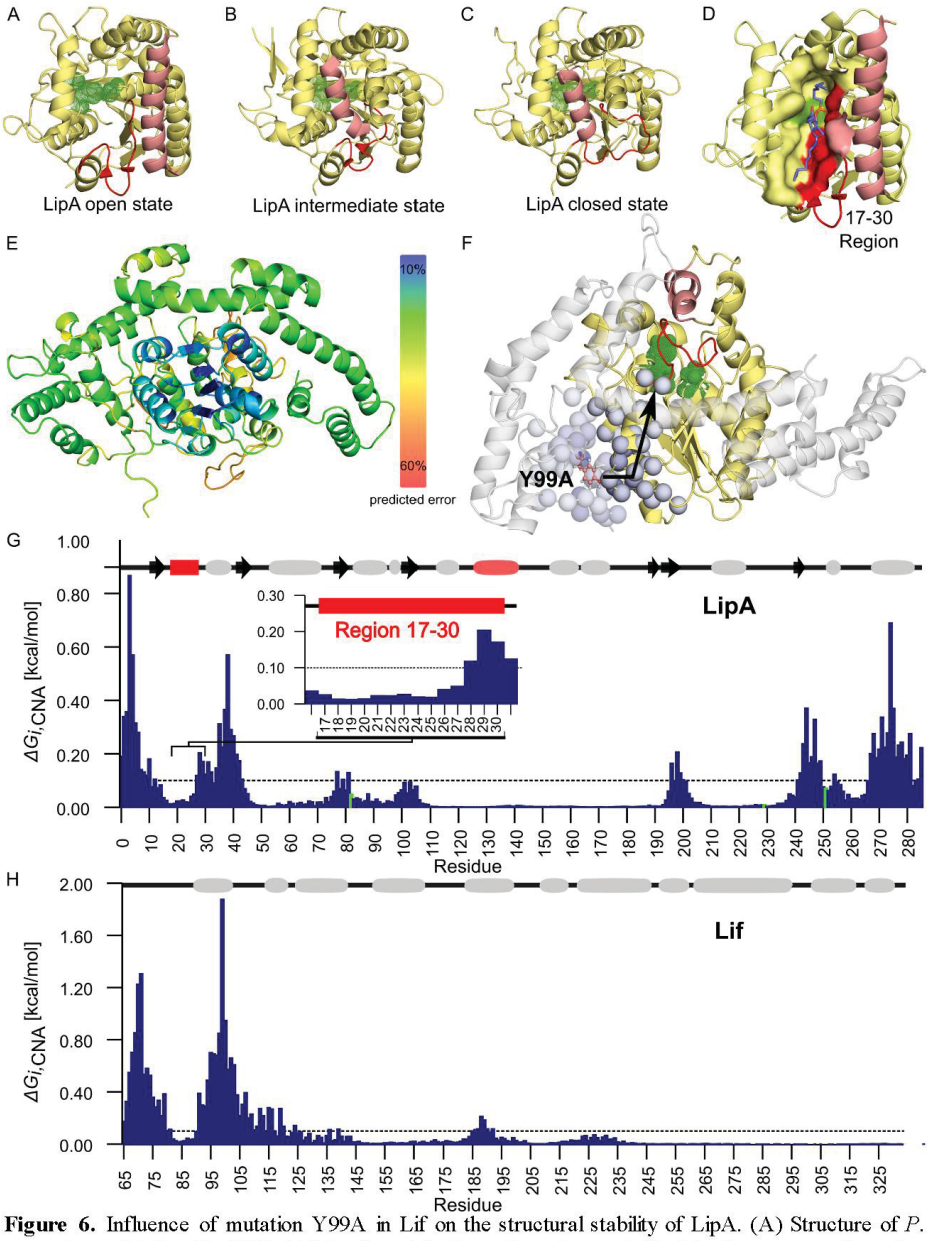


Figure 6. Influence of mutation Y99A in Lif on the structural stability of LipA. (A) Structure of *P. aeruginosa* LipA with OCP inhibitor bound in the active site crystallized in the open conformation (PDB code 1EX9(23)), in which the helix α5 (salmon) is moved away from the active site (catalytic triad residues S82, H251 and D229 shown in green). In this conformation, the active site is accessible for the substrate and LipA is enzymatically active. A short two-stranded β-sheet close to the active site is formed by residues 17-30 (red). (B) *B. glumae* lipase from the crystal structure of the Lif:LipA complex (PDB code 2ES4(10)). The lipase shows a two-stranded β-sheet (red), a characteristic feature

of the open (active) conformation, nevertheless helix α5 (salmon) adopts a closed (inactive) conformation. This suggests a foldase-induced formation of a two-stranded β-sheet during activation of the lipase. (C) Crystal structure of *B. glumae* lipase crystallized in the closed conformation (PDB code 1QGE(49)) with helix α5 (salmon) covering the active site (residues as in panel A). In this conformation, a two-stranded β-sheet close to the active site is not formed. Residues 17-30 of *B. glumae* lipase, forming a two-stranded β-sheet, are indicated red. (D) Crystal structure of active *P. aeruginosa* LipA with inhibitor OCP bound in the active site (PDB code 1EX9). Region of residues 17-30 forms part of the active site (red), required for the binding of the ligand. (E) Homology model of the *P. aeruginosa* Lif:LipA complex based on the structure of the *B. glumae* foldase-lipase complex (PDB code 2ES4(10)) used as a template. The coloring indicates the model quality assessment by TopScore,(*24, 25*) with bluish colors representing less than 10 % structural error. (F) CNA was applied on an ensemble of structures of the Lif:LipA complex generated from 10 independent MD simulations. Residues with $\Delta G_{i,\mathrm{CNA}}$ above a threshold of 0.1 kcal mol^{-1} are depicted as spheres on the Lif:LipA complex structure. Blue colors reflect predicted $\Delta G_{i,\mathrm{CNA}}$ values; the larger the value, the darker is the color. The black arrow indicates how the perturbation by Y99A mutation of Lif (pink, ball-and-stick representation) influences residues in LipA. Due to the decrease in the stability of the surrounding region of residues 17-30 in LipA, we speculate that the conformational changes required for the intermediate state of LipA on the way of activation is hampered upon $\mathrm{Lif}_{\mathrm{Y99A}}$ mutation. The color code for helix α5, residues 17-30 and the active site is as in panel A. (G) The histogram shows the per-residue $\Delta G_{i,\mathrm{CNA}}$ for LipA. The dashed line at 0.1 kcal mol^{-1} indicates the threshold above which residues are considered perturbed. The standard error of the mean is < 0.05 kcal mol^{-1} for all residues. (H) Per-residue $\Delta G_{i,\mathrm{CNA}}$ shown for Lif, with the same threshold. The standard error of the mean is < 0.05 kcal mol^{-1} for all residues.

SUPPORTING INFORMATION

Structural and dynamic insights revealing how lipase binding domain MD1 of *Pseudomonas aeruginosa* foldase affects lipase activation

Aldino Viegas[1,$,#], Peter Dollinger[2,$], Neha Verma[3,$], Jakub Kubiak[5], Thibault Vennet[1,4], Claus A.M. Seidel[5], Holger Gohlke[3,4,6,*], Manuel Etzkorn[1,4,7*], Filip Kovacic[2], and Karl-Erich Jaeger[2,8,*]

Running title: **Structural insights into foldase-mediated lipase activation**

From the [1]Institute of Physical Biology, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; [2]Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany; [3]Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; [4]Institute for Complex Systems ICS-6: Structural Biochemistry, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany; [5]Institute of Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; [6]John von Neumann Institute for Computing (NIC) and Jülich Supercomputing Centre (JSC), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany; [7]JuStruct: Jülich Center for Structural Biology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany; [8]Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

[#]Present address: Departamento de Química, Faculdade de Ciências e Tecnologia UCIBIO-REQUIMTE, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

*To whom correspondence should be addressed: Karl-Erich Jaeger: Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, D-52425, Jülich, Wilhelm Johnen Str. 1, Karl-Erich.jaeger@fz-juelich.de; Tel. (+49)02461616966; Fax. (+49)02461612490. Manuel Etzkorn: Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, D-40225, Düsseldorf, Universitätsstr. 1; manuel.etzkorn@hhu.de; Tel. (+49)2118112023; Fax. (+49)2118113847. Holger Gohlke: Institute of Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, D-40225, Düsseldorf, Universitätsstr. 1; gohlke@uni-duesseldorf.de; Tel. (+49) 2118113662; Fax. (+49)2118113847.

[$]These authors contributed equally to this work.

Supplementary Content

Supplementary Figures

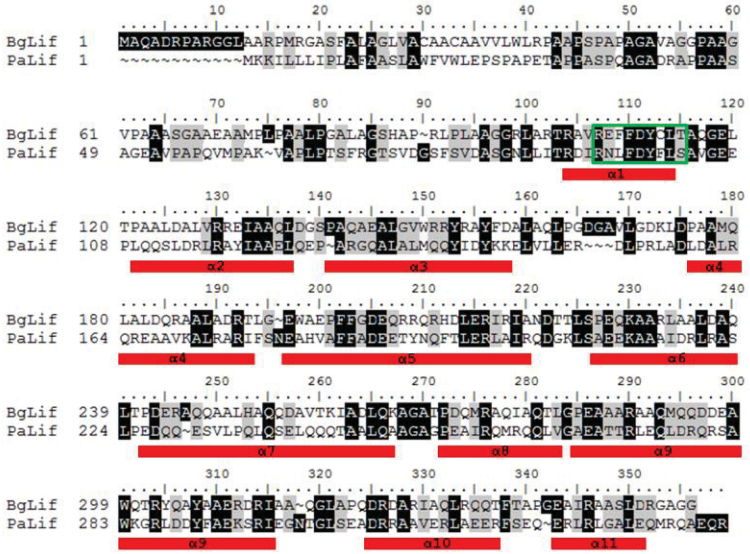


Figure S1 Sequence alignment of foldases from *B. glumae* and *P. aeruginosa*. BgLif and PaLif share 39% identical and 52% similar amino acids, shown on a black and gray background, respectively. Sequence similarity of *P. aeruginosa* and *B. glumae* TMD, VLD, MD1, EHD and MD2 is 31%, 44%, 52%, 46% and 48%, respectively. The red bars underneath the alignment indicate α-helices of the experimentally determined BgLif structure. The conserved RxxFDY(F/C)L(S/T)A foldase motif is indicated with a green frame. Numbers in front of the sequences indicate amino acid positions in the sequences.

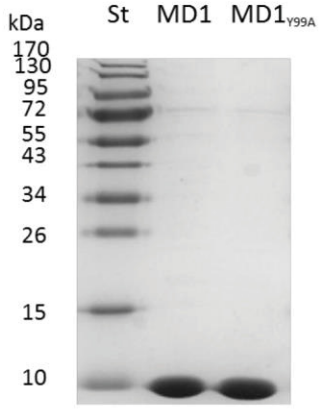


Figure S2 SDS-PAGE analysis of purified *P. aeruginosa* MD1 and MD1$_{Y99A}$ used for solution NMR analysis. The gel (16 %) was stained with Coomassie Brilliant Blue G250. Molecular weights of standard proteins (St) are indicated on the left-hand side.

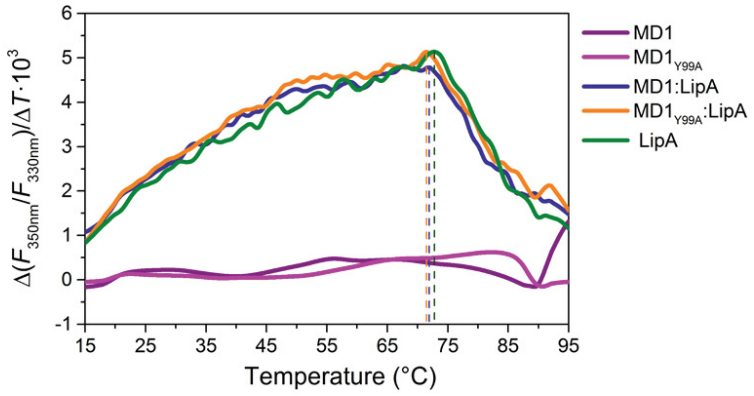


Figure S3 Melting curves of MD1, $MD1_{Y99A}$, pre-active LipA alone and after incubation with MD1 and $MD1_{Y99A}$ obtained by fluorescence measurement with nanoDSF. Vertical dashed lines indicate melting temperatures of LipA not affected significantly by incubation with MD1 or $MD1_{Y99A}$.

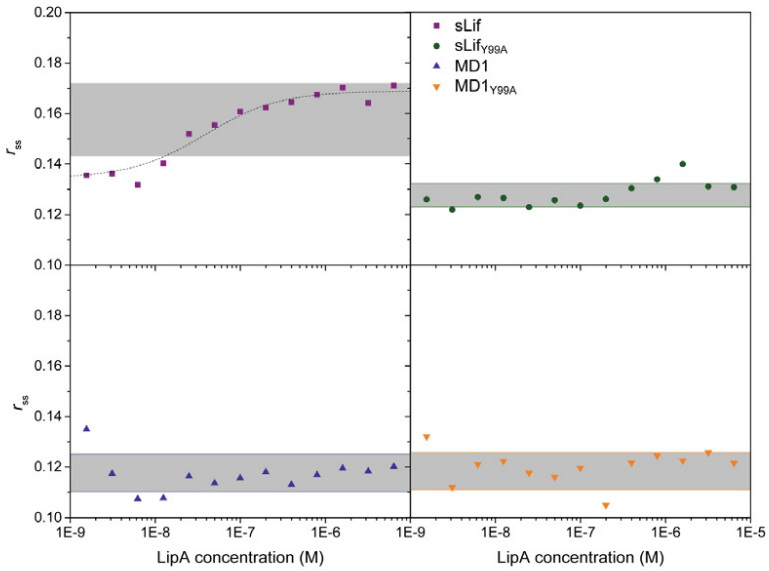


Figure S4 Steady-state fluorescence anisotropy $r_{steady\text{-}state}$ titration of sLif, $sLif_{Y99A}$, MD1 and $MD1_{Y99A}$ labelled with BDP FL NHS ester with different concentrations of LipA. Gray bar indicates average steady-state anisotropy ± the standard deviation. Steady-state anisotropy is a good marker of *s*Lif binding to LipA. However in case of $sLif_{Y99A}$, MD1 and $MD1_{Y99A}$ there is no observable change in this parameter at LipA concentrations of up to 10 μM. Steady-state anisotropy is a product of fluorescence decay and fluorescence anisotropy decay $r_{steady-state} = \int_0^\infty F(t_c)r(t_c)dt_c \,/\int_0^\infty F(t_c)dt_c$, which is sensitive i.a. to the dye flexibility and size of the rotating molecule. Stable steady-state anisotropy marks the lack of formation of complex between labelled molecules and LipA or change in other parameters influencing $r_{steady\text{-}state}$.

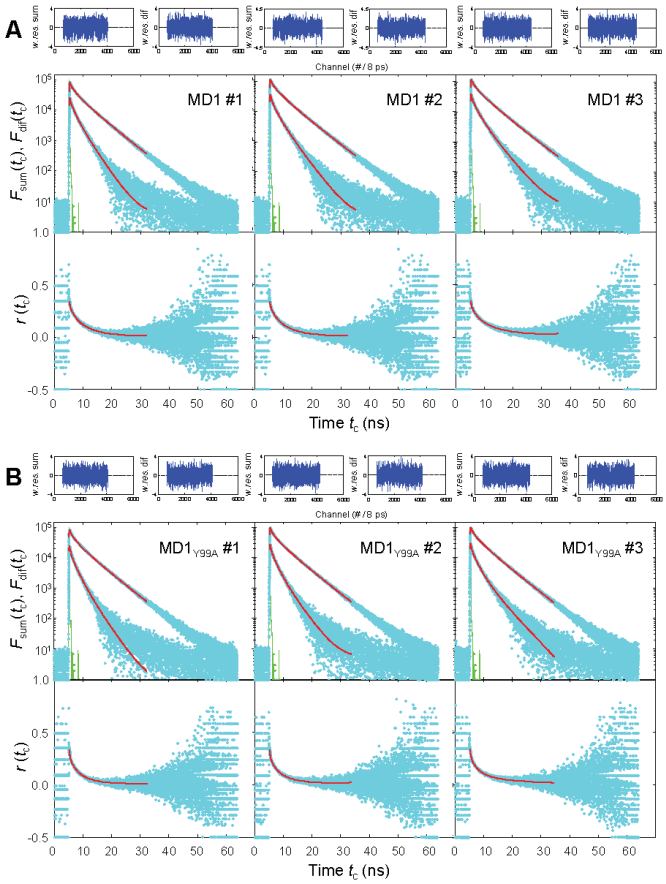


Figure S5. Interaction of fluorescently labeled MD1 (A) and MD1$_{Y99A}$ (B) with LipA: condition #1 - 0.1 µM MD1/MD1$_{Y99A}$; condition #2 - 0.1 mM MD1/MD1$_{Y99A}$ (NMR conditions); and condition #3 - 0.1 mM MD1/MD1$_{Y99A}$ and 0.4 mM LipA (NMR conditions). Joint time-resolved fluorescence anisotropy decay fits (sum and difference: ($F_{sum}(t_c) = F(t_c)$, $F_{dif}(t_c) = F(t_c) \cdot r(t_c)$). Fit parameters are collected in supplementary **Table S3**. MD1 and MD1$_{Y99A}$ show similar hydrodynamic properties and both experience slower rotation at 0.1 mM concentrations (condition #2), which canbe explained by increased viscosity (factor 1.2 compared to condition #1). Assuming global rotational correlation time of MD1:LipA complex of ρ_{global} = 32 ns, a fit of fluorescence anisotropy decay requires approx. 20% of such complex fraction in case of both MD1 and MD1$_{Y99A}$ in presence of 0.4 mM LipA (condition #3).

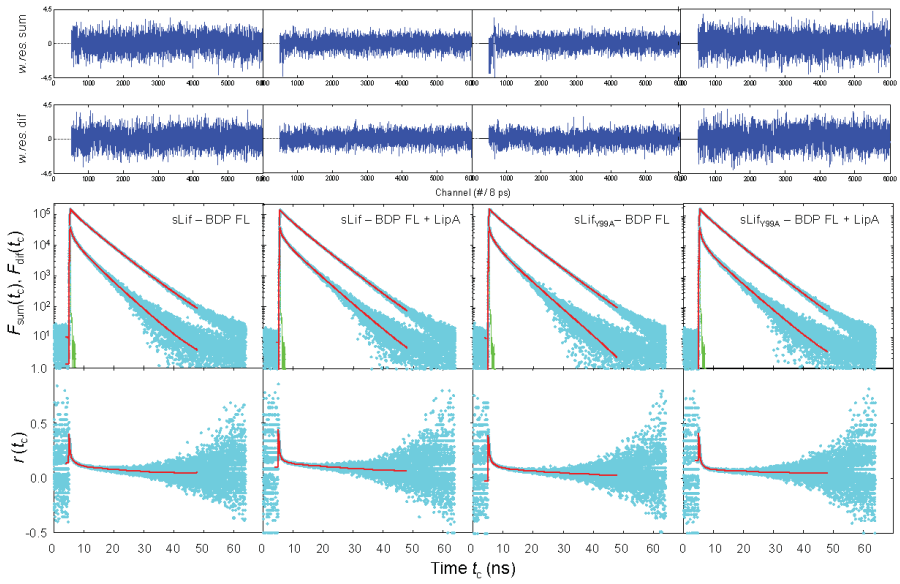


Figure S6 Time-resolved fluorescence anisotropy decay fits (sum and difference fit: $F_{sum}(t_c) = F(t_c)$, $F_{dif}(t_c) = F(t_c) \cdot r(t_c)$) of *s*Lif and *s*Lif$_{Y99A}$ labelled with BDP FL NHS ester in absence and presence of 40 µM sLipA. Amplitude of global rotational correlation time (r_3) is similar for *s*Lif and *s*Lif:LipA complex, while global rotational correlation time ρ_3 changest from approx. 33 ns to 50 ns. In case of *s*Lif$_{Y99A}$ ρ_3 shows similar behaviour but amplitude of global rotation (r_3) decreases significantly from 0.10 to 0.06. Fit parameters are collected in **Table S4**.

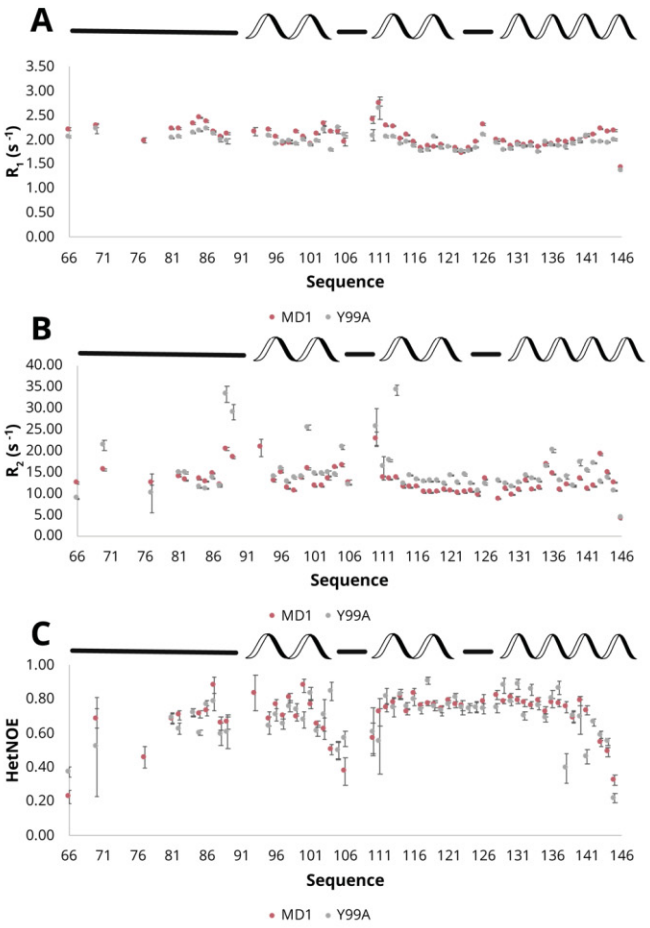


Figure S7 ^{15}N backbone relaxation data R_1, R_2 and heteronuclear {^{1}H}-^{15}N NOE (HetNOE) measured for MD1 (red) and MD1$_{Y99A}$ (grey) are shown in A, B, C, respectively. Secondary structure elements are illustrated on top of the figure. Data was acquired using ^{15}N-labelled samples at a concentration of 650 and 600 µM for MD1 and MD1$_{Y99A}$, respectively. The solutions were prepared either in Tris-Glycine buffer pH 9 containing 10% (v/v) D_2O, 0.01% sodium azide and 100 µM DSS or in 20 mM sodium phosphate buffer pH 7.4 containing 10% (v/v) D_2O, 0.01% sodium azide and 100 µM DSS for MD1 and MD1$_{Y99A}$, respectively. All data were collected in a Bruker Avance III HD$^+$ 600 MHz spectrometer, at 35°C.

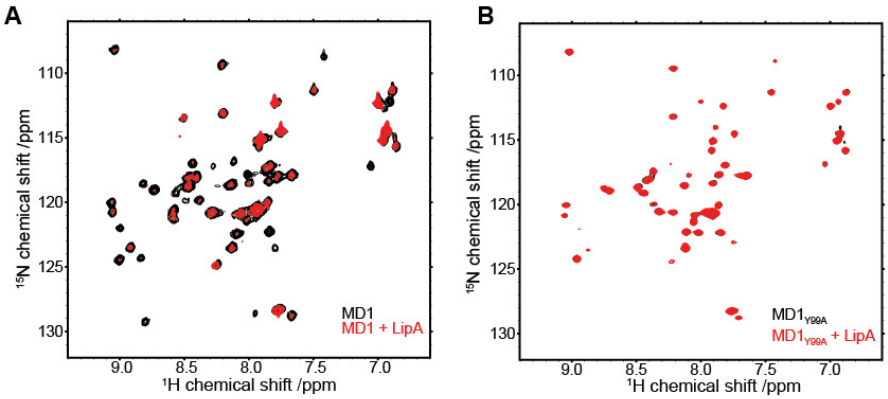


Figure S8 ^{1}H-^{15}N HSQC spectra of isotope labelled variants of MD1 in the absence (black) and presence of 3-fold molar excess of unlabeled LipA. (A and B) ^{1}H,^{15}N-HSQC spectra of MD1 and MD1$_{Y99A}$, respectively, in the presence (black) and absence of 3-fold molar excess of LipA (red). Note that the observed decrease in NMR signal intensity upon MD1-LipA interaction indicates binding in the NMR intermediate to slow exchange regime and that less effected peaks originate from nuclei located in the amino acids side chains or the protein's flexible regions. The relaxation properties of these positions should be less affected by the expected change in particle size upon interaction with the LipA. Overall, our data, therefore, do not allow to map the binding interface reliably.

Supplementary Tables

Table S1 Steady-state fluorescence anisotropy $r_{\text{steady-state}}$ and associated fraction of complex x_{complex} of sLif labelled with BDP FL NHS ester with LipA at different concentrations of LipA.

LipA [µM]	r_{ss}	x_{complex}	LipA [µM]	r_{ss}	x_{complex}
0.000	0.101	0.084	0.002	0.108	0.348
0.000	0.097	-0.023	0.002	0.105	0.252
0.000	0.096	-0.062	0.005	0.107	0.329
0.001	0.105	0.257	0.015	0.106	0.272
0.002	0.098	0.011	0.044	0.116	0.630
0.004	0.117	0.666	0.13	0.121	0.818
0.008	0.097	-0.048	0.40	0.129	1.085
0.016	0.107	0.309	1.20	0.120	0.756
0.032	0.105	0.245	3.59	0.126	0.996
0.063	0.114	0.542	0.002	0.106	0.279
0.13	0.115	0.588	0.005	0.105	0.259
0.25	0.123	0.860	0.015	0.103	0.171
0.50	0.128	1.039	0.044	0.113	0.530
1.01	0.122	0.855	0.13	0.120	0.785
2.02	0.133	1.228	0.40	0.124	0.919
4.04	0.122	0.829	1.20	0.126	0.984
8.08	0.126	0.996	3.59	0.134	1.272
0.002	0.100	0.054	10.8	0.122	0.839

Table S2 Average translational diffusion time $\langle t_{trans} \rangle$ (eq. 3) and associated fraction of complex $x_{complex}$ of sLif$_{Y99A}$ labelled with BDP FL NHS ester with LipA at different concentrations of LipA.

LipA [μM]	$\langle t_{trans} \rangle$ [ms]	$x_{complex}$	LipA [μM]	$\langle t_{trans} \rangle$ [ms]	$x_{complex}$
0.0016	0.75	0.219	1.2	0.92	1.058
0.0016	0.73	0.102	3.6	0.84	0.678
0.0049	0.72	0.097	10.8	0.93	1.115
0.015	0.65	-0.290	32.3	0.87	0.832
0.044	0.81	0.542	0.0005	0.72	0.051
0.13	0.80	0.491	0.0005	0.74	0.158
0.40	0.95	1.227	0.0005	0.71	0.006
1.2	0.91	1.032	0.0010	0.66	-0.228
3.6	0.88	0.890	0.0020	0.72	0.082
10.8	0.91	1.029	0.0039	0.66	-0.218
32.3	0.91	1.041	0.0079	0.66	-0.222
0.0016	0.75	0.212	0.016	0.77	0.315
0.0049	0.77	0.330	0.032	0.69	-0.082
0.015	0.70	-0.048	0.063	0.83	0.598
0.044	0.82	0.583	0.126	0.85	0.706
0.13	0.79	0.427	0.252	0.80	0.474
0.40	0.86	0.766	0.505	0.89	0.899

Table S3 Fluorescence anisotropy decay $r(t_c)$ fit parameters from the model function given by eq. S4b, of MD1 and MD1$_{Y99A}$ labelled with BDP FL NHS ester, with different concentrations of unlabelled MD1/MD1$_{Y99A}$ and LipA. Condition #1 - 0.1 μM MD1/MD1$_{Y99A}$. Condition #2 - 0.1 mM MD1/MD1$_{Y99A}$ (NMR conditions). Condition #3 - 0.1 mM MD1/MD1$_{Y99A}$ and 0.4 mM LipA (NMR conditions). MD1 and MD1$_{Y99A}$ show similar hydrodynamic properties and both experience slower rotation at 0.1 mM concentrations (condition #2), which canbe explained by increased viscosity (factor 1.2 compared to condition #1). Assuming global rotational correlation time of MD1:LipA complex of ρ_{global} = 32 ns, a fit of fluorescence anisotropy decay requires approx. 20% of such complex fraction in case of both MD1 and MD1$_{Y99A}$ in presence of 0.4 mM LipA (condition #3). Fits are shown in Fig. S5.

	Condition	r_1	ρ_1 [ns]	r_2	ρ_2 [ns]	r_3	ρ_3 [ns]	r_4	ρ_4 [ns]	r_{ss}	χ_r^2
MD1	#1					0.178	7.8[a]	0	0	0.162	
	#2	0.04	0.32	0.11	2.87	0.178	9.2	0	0	0.170	1.007
	#3					0.158	9.5[b]	0.024	32[c]	0.177	
MD1$_{Y99}$	#1					0.102	7.8[a]	0	0	0.129	
	#2	0.06	0.29	0.15	2.84	0.102	9.5	0	0	0.136	1.013
	#3					0.072	9.5[b]	0.039	32[c]	0.147	

[a]Global rotational correlation time estimated with HYDROPRO software [1] for MD1/MD1$_{Y99A}$ PDB code 5OVM at 20°C (viscosity 1.002 mPa·s).

[b]Global rotational correlation time of MD1/MD1$_{Y99A}$ in presence of 0.1 mM unlabelled MD1/MD1$_{Y99A}$ (NMR condition). Global rotation is approx. factor 1.2 slower, compared to sub-micromolar concentrations used for fluorescence measuremnets in condition #1, because of larger viscosity

[c]Global rotational correlation time of MD1/MD1$_{Y99A}$ in hypothetical complex with LipA estimated with HYDROPRO software [1] for PDB code 2ES4 with *bg*Lif replaced with MD1 (PDB code 5OVM, 20°C, 1.002 mPa·s, ρ_{global} = 26 ns) multiplied by factor 1.2.

Table S4 Fluorescence anisotropy decay $r(t_c)$ fit parameters.
A. sLif and $sLif_{Y99A}$ labelled with BDP FL NHS ester, in absence and presence of 40 μM LipA. Fluorescence decay $F(t_c)$ model parameters from the model function given by eq. S4a.

	x_1	τ_1 [ns]	x_2	τ_2 [ns]	x_3	τ_3 [ns]	a_3	τ_4 [ns]	$\langle\tau\rangle_x$ [ns]	$\langle\tau\rangle_F$ [ns]
sLif	0.471	6.03	0.390	3.67	0.090	1.16	0.048	0.249	4.56	5.27
+ LipA (40 μM)	0.600	5.89	0.277	3.43	0.76	0.78	0.046	0.130	4.55	5.30
$sLif_{Y99A}$	0.593	6.02	0.300	3.80	0.060	1.28	0.046	0.387	4.81	5.40
+ LipA (40 μM)	0.514	5.92	0.365	3.80	0.072	1.29	0.049	0.279	4.54	5.16

B. sLif and $sLif_{Y99A}$ labelled with BDP FL NHS ester, in absence and presence of 40 μM mM LipA. Fluorescence anisotropy decay $r(t_c)$ model parameters from the model function given by eq. S4b. Amplitude of global rotational correlation time (r_3) is similar for sLif and sLif:LipA complex, while global rotational correlation time ρ_3 changest from approx. 33 ns to 50 ns. In case of $sLif_{Y99A}$ ρ_3 shows similar behaviour but amplitude of global rotation (r_3) decreases significantly from 0.10 to 0.06. Fits are shown in **Figure** 6.

	r_1	ρ_1 [ns]	r_2	ρ_2 [ns]	r_3	ρ_3 [ns]	r_{ss}	χ_r 2
sLif	0.13	0.21	0.09	1.72	0.12	33	0.129	1.017
+ LipA (40 μM)	0.12	0.20	0.07	1.48	0.14	50	0.148	1.008
$sLif_{Y99A}$	0.13	0.22	0.09	1.89	0.10	33	0.116	1.085
+ LipA (40 μM)	0.15	0.20	0.09	1.42	0.08	50	0.101	1.028

Table S5 pFCS fit parameters for *s*Lif and $sLif_{Y99A}$ labeled with BDP FL from the model function $G(t_c)$ given by eq. S2, common fit parameters for 3 samples indicate a global optimization of these parameters in a free fit. Fits are shown in Fig. 3D.

Parameters	sLif	sLif:LipA $G_{p1,p2}(t_c)$	$sLif_{Y99A}$:LipA
$\chi 2$	24.82	12.45	17.98
N	1.56	1.37	1.07
t_{trans} [μs]	308	320	317
(z_0/w_0)		10.7	
b_1		0.136	
t_{b1} [μs]		83.7	
b_2		0.141	
t_{b2} [μs]		11.9	
b_3		0.112	
t_{b3} [μs]		1.4	
a	0.955	0.930	0.882
t_a [ns]		4.40	
b_{rot}	0.520	0.639	0.509
ρ_{global} [ns]	**33.2**	**49.8**	**49.0**
$C^{\#}$		0.0889	
$S^{\#}$		0.3	

fixed spherical rotator parameters, Kask *et al.* [2]

Table S6 List of used bacterial strains, plasmids and oligonucleotides.

Strain	**Genotype**	**Reference**
E. coli DH5α	F– Φ80**lac**ZΔM15 Δ(**lac**ZYA-**arg**F) U169 **rec**A1 **end**A1 **hsd**R17 (rK–, mK+) **pho**A **sup**E44 λ– **thi**-1 **gyr**A96 **rel**A1	Hanahan (1983) [3]
E. coli BL21 (DE3)	*E. coli* B dcm ompT hsdS($r_B^-m_B^-$) gal, λDE3	Studier *et al.*(1986) [4]
Plasmid	**Features**	**Reference**
pLipA-SS	pET22b (NdeI, BamHI, 916 bp; *lipA Δ1-78 bp* coding for signal sequence, Ser1Met, under P_{T7} control), Amp^r	Hausmann *et al.* (2008) [5]
pEHTHis19	pET19b (NdeI, BamHI, 985 bp; lipH Δ1-60bp coding for TMD, substitution 84 bp His_{10}-tag and soluble linker, under P_{T7} control), Amp^r	Hausmann *et al.* (2008) [5]
pET-MD1	pET19b carrying the gene encoding MGH_6 sequence in the front of MD1 (amino acids 66-146)	this work
pET-MD1$_{Y99A}$	Y99A mutation inserted into pET-MD1 plasmid	this work
pET-*s*Lif$_{Y99A}$	Y99A mutation inserted into pEHTHis19 plasmid	this work
Oligonucleotide	**Sequence**	
Lif_dLinkVD_fw	GGCCATCACCATCACCATCACCTGCCAACCTCCTTCAGGG	
Lif_dLinkVD_rv	CAGGTGATGGTGATGGTGATGGCCCATGGTATATCTCC	
Lif_Y99A_fw	ACATCCGCAACCTGTTCGACGCCTTCCTCAGCGCCGTCGGCG	
LipH_Y99_rv	GTCGAACAGGTTGCGGATGTCGC	
LipH_backbone_fw	TGACCGGCACGGAAACGC	
LipH_MD1_rv	TTTCCGTGCCGGTCACAGTTCCTTCTTGTAGTCGATG	

Table S7 Acquisition parameters of the spectra used for MD1 resonance assignment and structure calculation.

	Number of points			Spectral width (ppm)			Central frequency (ppm)			NS
Backbone assignment	F3	F2	F1	F3	F2	F1	F3	F2	F1	
2D										
$^{1}H,^{15}N$-HSQC	-	2048	256	-	12	30	-	4.695	119.5	64
$^{1}H,^{13}C$-HSQC	-	2048	512	-	12	75	-	4.695	39	64
3D										
HNCO	2048	40	128	12	36	22	4.706	119.5	176	8
HN(CA)CO	2048	40	128	12	36	22	4.706	119.5	176	16
HN(CO)CACB	2048	40	128	12	36	75	4.706	119.5	39	16
HNCACB*	2048	40	128	12	30	75	4.695	119.5	39	128
Side chain assignment										
(H)CCH-TOCSY	2048	64	128	12	75	75	4.706	39	39	16
NOE measurement										
3D- $^{1}H,^{15}N$-NOESY-HSQC (mix = 120 ms)	2048	40	120	12	30	12	4.695	119.5	4.695	32
3D- $^{1}H,^{13}C$-NOESY-HSQC (mix = 200 ms)	2048	40	120	12	75	12	4.695	39	4.695	32

* This experiment was recorded using non-uniformly sampling (NUS) of NMR data with 20 % sampling density. The NUS schedule was generated with the Poisson gap sampling method and the spectrum was subsequently reconstructed using hmsIST [6,7] and processed with NMRpipe [8]). All other spectra were processed with Topspin3.5 (Bruker BioSpin).

Table S8 Acquisition parameters of the spectra used for $MD1_{Y99A}$ resonance assignment and structure calculation.

	Number of points			Spectral width (ppm)			Central frequency (ppm)			NS
Backbone assignment	F3	F2	F1	F3	F2	F1	F3	F2	F1	
2D										
$^{1}H,^{15}N$-HSQC	-	2048	256	-	12	30	-	4.695	119.5	64
$^{1}H,^{13}C$-HSQC	-	2048	512	-	12	75	-	4.695	39	64
3D										
HN(CA)CO	2048	40	128	12	30	22	4.702	119.5	176	8
HNCACB	2048	40	128	12	30	75	4.697	119.5	39	16
Side chain assignment										
(H)CCH-TOCSY	2048	56	120	12	75	75	4.697	39	39	16
NOE measurement										
3D-$^{1}H,^{15}N$-NOESY-HSQC (mix = 250 ms)	2048	40	128	12	35	12	4.702	119.5	4.702	32
3D-$^{1}H,^{15}N$-NOESY-HSQC (mix = 120 ms)	2048	40	128	12	30	12	4.695	119.5	4.695	28
3D-$^{1}H,^{13}C$-NOESY-HSQC (mix = 200 ms)	2048	40	128	12	75	12	4.695	39	4.695	28

Supplementary Methods

Section S1. pFCS analysis

Polarization-resolved fluorescence correlation spectroscopy (pFCS), in which fluorescence intensity fluctuations under constant excitation are measured, is also able to resolve molecular rotational motion. Auto- and cross-correlations between different polarization channels were obtained were obtained according to Felekyan *et al.* [9]

$$G_{s1,s2}(t_c) = 1 + \frac{\langle \delta F_{s1}(t) \delta F_{s2}(t+t_c) \rangle}{\langle F_{s1}(t) \rangle \langle F_{s2}(t) \rangle}$$

$$G_{p1,p2}(t_c) = 1 + \frac{\langle \delta F_{p1}(t) \delta F_{p2}(t+t_c) \rangle}{\langle F_{p1}(t) \rangle \langle F_{p2}(t) \rangle}$$

$$G_{s,p}(t_c) = 1 + \frac{\langle \delta F_{s}(t) \delta F_{p}(t+t_c) \rangle}{\langle F_{s}(t) \rangle \langle F_{p}(t) \rangle}$$

$$G_{p,s}(t_c) = 1 + \frac{\langle \delta F_p(t)\, \delta F_s(t+t_c) \rangle}{\langle F_s(t) \rangle \langle F_p(t) \rangle} \quad \text{(S1)}$$

with the fluorescence fluctuations, $\delta F(t) = F(t) - \langle F(t) \rangle$. The registered photon events were analyzed by employing a custom designed software package for multiparameter fluorescence spectroscopy, full correlation and multiparameter fluorescence imaging [9]. The applied factorized fitting function (eq S4) models translational diffusion in a 3D-Gaussian volume element $G_{trans}(t_c)$, up to three temporary dark states $G_b(t_c)$, rotational diffusion of a spherical rotator $G_{rot}(t_c)$ and photon antibunching $G_a(t_c)$:

$$G(t_c) = 1 + \frac{1}{N} G_a(t_c) \times G_b(t_c) \times G_{rot}(t_c) \times G_{trans}(t_c) \quad \text{(S2)}$$

with

$$G_{trans}(t_c) = \left(1 + \frac{t_c}{t_{trans}}\right)^{-1} \left(1 + \left(\frac{\omega_0}{z_0}\right)^2 \times \frac{t_c}{t_{trans}}\right)^{-\frac{1}{2}} \quad \text{(S2a)}$$

$$G_b(t_c) = 1 - b_1 + b_1 e^{-t_c/t_{b1}} - b_2 + b_2 e^{-t_c/t_{b2}} - b_3 + b_3 e^{-t_c/t_{b3}} \quad \text{(S2b)}$$

$$G_a(t_c) = 1 - a e^{-t_c/t_a} \quad \text{(S2c)}$$

$$G_{rot}(t_c) = 1 + b_{rot}\left(\frac{1}{1+C} e^{-t_c/\rho_{global}} + \frac{C}{1+C} e^{-t_c/\left(S\rho_{global}\right)}\right) \quad \text{(S2d)}$$

Here, the observation volume is approximated by a 3D-Gaussian volume with $1/e^2$ radii in the lateral (ω_0) and axial direction (z_0), t_{trans} is the diffusion time, $b_{1,2,3}$ and $t_{b1,b2,b3}$ are amplitudes and times of the bunching terms, N is the particle number, a and t_a are the amplitude and time of the antibunching term, S and C characterize the rotation model (see Kask *et al.* [2]), b_{rot} and ρ_{global} are the amplitude and correlation time associated with rotational motion. As described above, factorization of the model function (eq S4) is based on the assumption of well-separated time scales for antibunching ($t_a \approx \tau_e$) and rotational correlation (ρ_{global}).

Section S2. Fluorescence anisotropy decay analysis

Polarization resolved fluorescence intensity decay histograms $F_p(t_c)$ and $F_s(t_c)$ were recorded. The fluorescence $F(t_c)$ and anisotropy decay $r(t_c)$ parameters were recovered by global fitting o the sum $F_{sum}(t_c)$ and difference $F_{dif}(t_c)$ histograms as previously described by Mockel *et al.* [10]

$$F_{sum}(t_c) = \left(F_p(t_c) + F_s(t_c)\right) = F(t_c) \quad \text{(S3a)}$$

$$F_{dif}(t_c) = \left(F_p(t_c) - F_s(t_c)\right) = F(t_c) \cdot r(t_c) \quad \text{(S3b)}$$

where fluorescence and anisotropy decays were modelled by a weighted sum of exponentials,

$$F(t_c) = \sum_{i=1}^{4} x_i \cdot \tau_i \text{ where } \sum x_i = 1 \quad \text{(S4a)}$$

$$\text{and } r(t_c) = \sum_{i=1}^{4} r_i \cdot \rho_i \quad \text{(S4b)}$$

with fluorescence lifetimes τ_i, their species fractions x_i, and with rotational correlation times ρ_i and their related amplitudes r_i. We checked that $\sum_{i=1}^{4} r_i \leq r_0$ with the fundamental anisotropy of BDP FL r_0=0.375.[11]

Order parameters of global rotational correlation time[12] were calculated according to [10]: $S^2 = r_3/r_0$.

Section S3. Determination of half-inactivation temperature.

Half-inactivation temperature T_{50}, i.e. the temperature at which LipA activity is reduced to 50% in a temperature-dependent lipase activity of *s*Lif and LipA generated by incubation of pre-active LipA (100 nM) with *s*Lif (250 nM) overnight at 4°C in TG buffer. Samples were then incubated at different temperatures (10 – 50°C) for 1 h followed by measurement of the remaining lipase activity with 2 nM LipA by *p*NPP-based lipase activity assay. Data were fitted according to:

$$A(T) = \frac{A_{max} - A_{min}}{1 + e^{(T - T_{50}/\partial T)}} + A_{min} \quad \text{(S5)}$$

where $A(T)$ is lipase activity at temperature T, A_{max} is maximum enzyme activity (here 100%) and A_{min} is minimum enzyme activity (here 0%), δT is a rate of activity loss. Half-inactivation temperature T_{50} was determined to 29.0 ± 0.2°C and 2.1 ± 0.2°C^{-1}.

Supplementary References

1. Ortega, A., Amoros, D., and de la Torre, J. G. (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys J* **101**, 892-898
2. Kask, P., Piksarv, P., Pooga, M., Mets, U., and Lippmaa, E. (1989) Separation of the rotational contribution in fluorescence correlation experiments. *Biophys J* **55**, 213-220
3. Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* **166**, 557-580
4. Studier, F. W., and Moffatt, B. A. (1986) Use of bacteriophage-T7 RNA-polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology* **189**, 113-130
5. Hausmann, S., Wilhelm, S., Jaeger, K. E., and Rosenau, F. (2008) Mutations towards enantioselectivity adversely affect secretion of *Pseudomonas aeruginosa* lipase. *FEMS Microbiol Lett* **282**, 65-72
6. Hyberts, S. G., Milbradt, A. G., Wagner, A. B., Arthanari, H., and Wagner, G. (2012) Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *Journal of Biomolecular Nmr* **52**, 315-327
7. Hyberts, S. G., Robson, S. A., and Wagner, G. (2013) Exploring signal-to-noise ratio and sensitivity in non-uniformly sampled multi-dimensional NMR spectra. *Journal of Biomolecular Nmr* **55**, 167-178
8. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) Nmrpipe - a multidimensional spectral processing system based on unix pipes. *Journal of Biomolecular Nmr* **6**, 277-293
9. Felekyan, S., Kuhnemuth, R., Kudryavtsev, V., Sandhagen, C., Becker, W., and Seidel, C. A. M. (2005) Full correlation from picoseconds to seconds by time-resolved and time-correlated single photon detection. *Rev Sci Instrum* **76**
10. Mockel, C., Kubiak, J., Schillinger, O., Kuhnemuth, R., Della Corte, D., Schroder, G. F., Willbold, D., Strodel, B., Seidel, C. A. M., and Neudecker, P. (2019) Integrated NMR, fluorescence, and molecular dynamics benchmark study of protein mechanics and hydrodynamics. *J Phys Chem B* **123**, 1453-1480
11. Karolin, J., Johansson, L. B. A., Strandberg, L., and Ny, T. (1994) Fluorescence and absorption spectroscopic properties of dipyrrometheneboron difluoride (Bodipy) derivatives in liquids, Lipid-Membranes, and Proteins, *J Am Chem Soc 116*, 7801-7806.
12. Lipari, G., and Szabo, A. (1980) Effect of librational motion on fluorescence depolarization and nuclear magnetic-resonance relaxation in macromolecules and membranes, *Biophys J 30*, 489-506.

12. ORIGINAL PUBLICATION III (Submitted 2019)

The membrane-integrated steric chaperone Lif facilitates active site opening of *Pseudomonas aeruginosa* lipase A

Neha Verma[1], Peter Dollinger[2], Filip Kovacic[2], Karl-Erich Jaeger[2,3], and Holger Gohlke[1,4] *

[1]Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

[2]Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, Forschungszentrum Jülich GmbH, 52426 Jülich, Germany

[3]Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52426 Jülich, Germany

[4]John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC) and Institute for Complex Systems - Structural Biochemistry (ICS-6), Forschungszentrum Jülich GmbH, 52426 Jülich, Germany

Running title: Lif facilitates active site opening in LipA

Keywords: protein folding, activation, potential of mean force, molecular dynamics simulations, rigidity analysis

*Address: Universitätstr. 1, 40225 Düsseldorf, Germany.
Phone: (+49) 211 81 13662, Fax: (+49) 211 81 13847
E-mail: gohlke@uni-duesseldorf.de

Abstract

Lipases are essential and widely used biocatalysts. Hence, the production of lipases requires a detailed understanding of the molecular mechanism of its folding and secretion. Lipase A from *Pseudomonas aeruginosa, Pa*LipA, constitutes a prominent example that has additional relevance because of its role as a virulence factor in many diseases. *Pa*LipA requires the assistance of a membrane-integrated steric chaperone, the lipase-specific foldase Lif, to achieve its enzymatically active state. However, the molecular mechanism of how Lif activates its cognate lipase has remained elusive. Here, we show by molecular dynamics simulations at the atomistic level and potential of mean force computations that Lif catalyzes the activation process of *Pa*LipA by structurally stabilizing an intermediate *Pa*LipA conformation, particularly a β-sheet in the region of residues 17-30, such that the opening of *Pa*LipA's lid domain is facilitated. This opening allows substrate access to *Pa*LipA's catalytic site. A surprising and so far not fully understood aspect of our study is that the open state of *Pa*LipA is unstable compared to the closed one according to our computational and *in vitro* biochemical results. We thus speculate that further interactions of *Pa*LipA with the Xcp secretion machinery and/or components of the extracellular matrix contribute to the remaining activity of secreted *Pa*LipA.

Introduction

P. aeruginosa lipase A (*Pa*LipA) requires the assistance of a membrane-integrated steric chaperone, the lipase-specific foldase Lif, to achieve its enzymatically active state. Here, we show by unbiased and biased molecular dynamics simulations at the atomistic level and potential of mean force computations that Lif catalyzes the activation process of *Pa*LipA by structurally stabilizing an intermediate *Pa*LipA conformation. *Pa*LipA is an important and widely used enzyme in synthetic applications because it catalyzes the hydrolysis and synthesis of a broad range of substrates.[1,2] Like other lipases, *Pa*LipA has a core structure comprised of the α/β-hydrolase fold, an active site with the catalytic triad consisting of Ser82, His251, and Asp229, and an oxyanion hole formed by Met16 and His83.[1] A characteristic feature of many lipases is "interfacial activation", which describes the fact that lipase activity increases in the presence of insoluble substrates that form an emulsion.[3] In the closed or inactive state, the active site of these lipases is covered by a lid, and this lid opens upon binding of the lipases to a hydrophobic interface.[4] *Pa*LipA possesses a lid formed by α-helix 5 but does not show "interfacial activation".[5] Rather, the production of enzymatically active *Pa*LipA is a complex process involving about 30 proteins for proper folding and extracellular secretion.[6] In a critical step, *Pa*LipA requires the assistance of an inner membrane-bound steric chaperone, the lipase-specific foldase Lif, for its conversion into an open conformation, which is also active.[7] Subsequently, *Pa*LipA is secreted to the extracellular medium *via* the type II secretion pathway using the Xcp machinery.[8] Without Lif, *Pa*LipA only folds to a near-native, but non-active state.[9,10]

Lif belongs to a small class of steric chaperones that act by lowering the energy barrier between a near-native state and an active state of the target protein.[9] Lif proteins catalyze the folding process by imprinting the essential steric (structural) information onto the target protein.[11] In that respect, steric chaperones differ from classical molecular chaperones, which indirectly increase the folding efficiency by preventing off-pathway interactions for newly synthesized proteins.[11] Most remarkably, the active state of enzymes that need specific steric chaperones is less or only marginally more stable than the inactive intermediate state.[12] However, the molecular mechanism of how these Lif proteins activate their cognate lipases has remained elusive.

A crystal structure of the complex of the homologous *Burkholderia glumae* lipase with its specific foldase (Lif:*Bg*LipA) (PDB code: 2ES4)[13] revealed that the core structure of the lipase

is virtually identical to that of the closed conformation of *B. glumae* LipA (*Bg*LipA$_c$) (PDB code: 1QGE)[14] and the open state of *Pa*LipA (*Pa*LipA$_o$) (PDB code: 1EX9)[1]. The circular dichroism spectrum of the *Bg*LipA$_c$ was found indistinguishable to that of the active *Bg*LipA, which indicates no major change in secondary structure upon activation.[9] Yet, major differences were found in two key regions: first, the lid domain (helix 5) and neighboring loops (residues 109-163) occlude the active site in *Bg*LipA$_c$ and the Lif:*Bg*LipA complex but are more distant from the active site in the *Pa*LipA$_o$ structure (Figure 1A). Furthermore, helix 5 is longer by ten residues at the N-terminal end (residue 125-148) in *Pa*LipA$_o$ than in *Bg*LipA$_c$ (residues 135 -148). Second, residues 17-30 form a partial β-sheet structure in the Lif:*Bg*LipA complex and *Pa*LipA$_o$, while they form a loop in *Bg*LipA$_c$. Residues 17-30 contribute to the formation of the active site surface in *Pa*LipA$_o$, and the partial β-sheet formation (residues 21-26) likely stabilizes neighboring loops (residues 17-20 and 27-32) (Figures 1B and 1C). Together, these observations led us to hypothesize that the foldase-bound *Bg*LipA is in an intermediate conformation where partial β-sheet formation has occurred but not yet lid opening, this conformation can be considered a "loaded spring" ready to change to the open conformation.

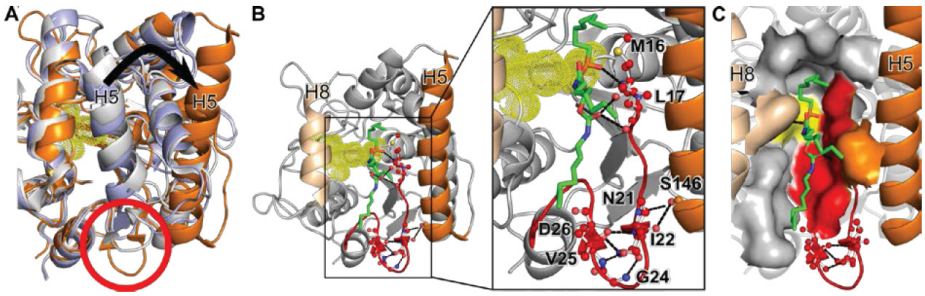


Figure 1. Structural superimposition of *B. glumae* lipase A (*Bg*LipA) and *P. aeruginosa* lipase A (*Pa*LipA) and schematic view of the active site. (A) Overlay of *Bg*LipA extracted from the foldase:lipase complex of *B. glumae* (gray, PDB code: 2ES4), closed *Bg*LipA (*Bg*LipA$_c$) (blue, PDB code: 1QGE), and *Pa*LipA in the open conformation (*Pa*LipA$_o$) (orange, PDB code: 1EX9). Extracted *Bg*LipA shows a partial β-sheet structure of region 17-30 residues, which is also present in the *Pa*LipA$_o$ structure but absent in the *Bg*LipA$_c$ structure (red circle). In contrast, helix 5 (H5) resembles the closed state by occluding the active site (yellow) in both conformations of *Bg*LipA but is moved away in *Pa*LipA$_o$, indicated by the black arrow. **(B)** *Pa*LipA$_o$ bound to an octyl-phosphinic acid 1,2-bis-octylcarbamoyloxy-ethyl ester (OCP) inhibitor (green sticks) is shown to visualize the active site (PDB code: 1EX9). H5 is shown in orange, helix 8 (H8) in wheat, and residues 17-30 in red. The catalytic triad is shown as yellow dots. The close-up view of the binding site (right inlay) shows that residues 17-30 form multiple intramolecular polar interactions in the region of partial β-sheet formation (residues 21-26). M16 forms

the oxyanion hole. **(C)** As in panel B but with a surface representation of the binding site to show that H5, H8, and residues 17-30 contribute to the formation of the active site surface.[1]

We probed this hypothesis by unbiased and biased molecular dynamics (MD) simulations followed by configurational free energy computations, complemented by *in vitro* biochemical experiments for validation. Our results indicate that Lif catalyzes the activation process of *Pa*LipA by structurally stabilizing the intermediate conformation, which facilitates the opening of the lid domain.

Methods

Generation of starting structures

The three-dimensional structure of *Pa*LipA in the closed conformation (*Pa*LipA$_c$) and complex with its foldase Lif (Lif:*Pa*LipA) is currently unknown. Thus, the homology model of *Pa*LipA$_c$ as well as the Lif:*Pa*LipA complex were constructed using the crystal structure of *B. glumae* lipase in complex with its foldase (PDB code: 2ES4) (sequence identity/similarity: 41%/73% for *Pa*LipA and 39%/52% for Lif) and the open *Pa*LipA (*Pa*LipA$_o$) (PDB code: 1EX9) as the templates. The Phyre2 web server[15] was used for homology modeling, followed by ten rounds of energy minimization with the GROMOS96 43B1 force field implemented in Swiss-PdbViewer.[16] The best model obtained was re-evaluated by using our in-house model quality assessment program TopScore.[17] The starting structure of *Pa*LipA$_o$ was obtained from the coordinates of the X-ray structure (PDB code: 1EX9).

Molecular dynamics simulations

All-atom MD simulations were performed with the Amber14 software package[18], using the ff14SB force field[19] as done previously by us.[20] *Pa*LipA$_c$, the Lif:*Pa*LipA complex, and *Pa*LipA$_o$ were placed in truncated octahedral periodic boxes of TIP3P water molecules, respectively.[21] The Particle Mesh Ewald (PME) method[22] was used to treat long-range electrostatic interactions and the SHAKE algorithm[23] to constrain the length of bonds to hydrogen atoms. A time step of 2 fs was used with a non-bonded cut-off of 8 Å. Initially, the starting structures were energy minimized by applying 50 steps of steepest descent, followed by 450 steps of conjugate gradient minimization. During the initial minimization, harmonic restraints with a force constant of 25 kcal mol^{-1} Å^{-2} were applied to the solute atoms and then reduced to 5 kcal mol^{-1} Å^{-2}. The systems were heated from 100 K to 300 K for thermalization by MD simulations in the canonical (NVT) ensemble, using the weak-coupling algorithm for

temperature control[24], carried out for 50 ps and using a force constant of 5 kcal mol^{-1} Å^{-2}. Afterward, MD simulations of 250 ps length were performed using isothermal-isobaric (NPT) ensemble MD simulations using the isotropic Berendsen barostat[24] with the same force constant in order to adjust the density of the system. Then, the force constant of the harmonic restraints was reduced to zero, and MD simulations in the NVT ensemble were carried out for 100 ps. Finally, six production MD simulations of 1 µs length each were performed for each of the three systems in the NVT ensemble using the weak-coupling algorithm for temperature control[24] with a coupling parameter $\tau = 1$ ps. To ensure the independence of the simulations, production runs were performed at temperatures of 300.0 K + T, where T was varied by 0.1 K from 0.0 to 0.5 K, respectively.[25]

Potential of mean force computations

For the potential of mean force (PMF) computations, the transition pathway of H5 between the open and closed conformations was taken from the unbiased MD simulations of *Pa*LipA$_o$. The start and end conformations of *Pa*LipA with closed and open lid were selected based on the distance between the centers of mass (D_{COM}) of the lid domain (H5, residues 125-148) and H8 (residues 210-222), which is 11.6 Å (minimum distance found during MD simulations) in the closed state and 20.6 Å in the open state. The closed state resembles the homology model of *Pa*LipA$_c$ (D_{COM} = 13.3 Å), and the crystal structure of *Pa*LipA$_o$ was taken as the open state.
The free energy profile of the opening of the active site was calculated for the *Pa*LipA structure and the Lif:*Pa*LipA complex, using umbrella sampling MD simulations in combination with the WHAM method.[26] D_{COM} was used as a reaction coordinate. Umbrella sampling MD simulations were performed along the reaction coordinate between 11.6 Å and 20.6 Å in intervals of 1 Å, applying a harmonic potential with a force constant of 2 kcal mol^{-1} Å^{-2} to tether the conformations to the respective reference point. This resulted in 10 umbrella sampling simulations per system, each 650 ns long. The first 50 ns were excluded from the subsequent WHAM analysis. The errors of the PMF profiles at the reference points were computed by applying the Monte Carlo bootstrapping analysis as implemented in WHAM using 400 resampling trials.

Analysis of trajectories

The unbiased MD trajectories were analyzed with the Amber module CPPTRAJ.[27] For each system, the average β-sheet propensity and D_{COM} were calculated; the former was calculated for residues 17-30 using the DSSP command. Additionally, the unbiased MD-generated conformations of *Pa*LipA$_c$, Lif:*Pa*LipA, and *Pa*LipA$_o$ were clustered with respect to D_{COM}. For the cluster analysis, the hierarchical agglomerative algorithm was used. A maximal distance between all members of two clusters (complete linkage) of 4 Å was used as ending criterion for the clustering. With these settings, we obtained a total of five clusters for each system. For the alignment of the structures onto the respective starting structures, root mean square fitting was done on the core residues (1-108 and 164-285) of *Pa*LipA for all systems.

Likewise, configurations *t* obtained by umbrella sampling in the windows corresponding to states I-III (see below for a definition of these states) were analyzed. To "unbias" these configurations, a weight w_t according to eqs 7 and 8 from ref.[28] was computed as done previously by us.[29] The reweighting is performed over the entire ensemble of each system and, then, w_t is normalized with respect to the sum of all w_t of each system. Finally, to identify interactions between residues in Lif and the key regions of *Pa*LipA, the average Cα-Cα distance matrix was calculated for Lif:*Pa*LipA over the six unbiased trajectories. An interaction is considered formed between respective two residues if the distance is < 10 Å, considering that the average length of the side chain is 3.5 Å for an amino acid.

Statistical analysis

Results for the β-sheet propensity of each residue and the Cα-Cα distance matrix computed from six unbiased MD simulations are expressed as arithmetic mean ± standard error of the mean (SEM). The overall SEM was calculated using the law of error propagation (eq. 1)

$$\mathrm{SEM}_{total} = \sqrt{\mathrm{SEM}_1^2 + \mathrm{SEM}_2^2 + \cdots + \mathrm{SEM}_6^2} \qquad \text{(eq. 1)},$$

where SEM_i is the SEM over each trajectory *i*. Following ref.[30], SEM_i was computed considering the decorrelation time of the examined variable. To analyze if averaged β-sheet propensities are statistically significantly different between systems, the Student's *t*-test[31] was applied, *p*-values < 0.05 and 0.001 are indicated as "*" and "**" in figures, respectively. The statistical analysis was performed using the R software.[32]

Cloning, protein production, and purification

The expression plasmid encoding *Pa*LipA and Lif was created by PCR using the Phusion High-Fidelity DNA polymerase (Thermo Fischer Scientific) in whole plasmid amplification designed for the SLIC method.[33] pLipA-SS and pEHTHis19 plasmids were used as templates for *Pa*LipA and Lif, respectively. For Lif, amino acids 1-65 were deleted using primers Lif_dLinkVD_fw and Lif_dLinkVD_rv. The expression plasmid for the *Pa*LipA variant with mutation S82A was created by whole plasmid PCR amplification with mutagenic oligonucleotide pair LipA_S82A_FW/LipA_S82A_RV designed for the SLIC method. *Pa*LipA, *Pa*LipA$_{S82A}$, and Lif were expressed in *E. coli* BL21 (DE3) using the T7-expression system with the respective plasmids. Lif was purified by immobilized metal affinity chromatography according to the modified protocol of Hausmann *et al.*[34] Cells expressing insoluble inclusion bodies of *Pa*LipA or *Pa*LipA$_{S82A}$ were suspended in Tris-HCl buffer (100 mM, pH 7) containing 5 mM EDTA and 1 mM Tris(2-carboxyethyl)phosphine (TCEP) and disrupted with a French press. Inclusion bodies were collected by centrifugation at 10,000 *g* for 10 min and suspended in the same buffer. Centrifugation and wash steps were repeated three times to obtain purified inclusion bodies. The purified inclusion bodies were suspended in a small amount of water and, afterward, 8 M urea (0.5 ml / 100 ml culture volume) was added. The inclusion bodies were incubated for 1 h at 37°C or until all inclusion bodies have been dissolved.

In vitro activation of *Pa*LipA with Lif.

Chemically denatured *Pa*LipA and *Pa*LipA$_{S82A}$ inclusion bodies were renatured by fast, at least 100-fold, dilution of the denaturant with ice cold 10 mM TG (5 mM TRIS, 5 mM glycine, pH 9) containing an equimolar amount of Lif followed by overnight incubation at 4°C.

Lipase activity assay

Para-nitrophenyl palmitate (*p*NPP, 1 mM) was used as a substrate in 10 mM TG buffer containing 1 mM $CaCl_2$ to determine lipolytic activities.[35] The release of *p*-nitrophenolate was monitored spectrophotometrically.

Protein stability determination by differential scanning fluorimetry

Lif:*Pa*LipA and Lif:*Pa*LipA$_{S82A}$ complexes prepared as described above were loaded into the measuring capillaries (Prometheus NT.Plex nanoDSF Grade Standard Capillary Chips) from a

384-well microtiterplate and were heated from 15°C to 95°C (heating rate of 0.2°C/min). The emission shift over temperature (F) was recorded at 330 nm and 350 nm using the Prometheus NT.Plex nanoDSF device (Nano Temper, Munich, Germany). The PR.ThermControl software provided by the company was used to calculate the ratio of $F_{350\ nm}$ and $F_{330\ nm}$ and its first derivative.

Constraint network analysis

To quantify a change in structural rigidity of *Pa*LipA upon binding to Lif, we employed a perturbation approach[36] using the Constrained Network Analysis (CNA) methodology[37], as described previously.[38] Briefly, CNA is a graph theory-based tool for rigidity analysis and has successfully been applied to a number of problems.[39-41] In a perturbation approach, the rigidity analysis is compared before (ground state) and after perturbing the constraint network by removing constraints of the residues of interest. For the perturbation analysis, first, an ensemble of network topologies was generated from MD snapshots of the Lif:*Pa*LipA complex, sampled at 2 ns intervals from the six unbiased MD simulations of 1 µs length each. Second, Lif residues forming interactions with the key regions in *Pa*LipA were identified from the average Cα-Cα distance matrix as described above. Third, for each of the identified residues *i* in Lif (residues 195-203, 213,217-220), the perturbation was performed, which resulted in a per-residue perturbation free energy $\Delta G_{i,\mathrm{CNA}}$ following a linear response approximation (eq. 2):

$$\Delta G_{i,\mathrm{CNA}} = \alpha(\langle E_{i,\mathrm{CNA}}^{perturbed}\rangle - \langle E_{i,\mathrm{CNA}}^{ground}\rangle) \quad (2)$$

α was set to 0.02 as in ref.[36]

Results

Structural dynamics of the lid of free *Pa*LipA and when bound to Lif

Initially, we aimed at analyzing the tendency of closed *Pa*LipA (*Pa*LipA$_c$) free and in complex with Lif (Lif:*Pa*LipA) to move towards the open state, and of free open *Pa*LipA (*Pa*LipA$_o$) towards the closed state, by unbiased MD simulations. Due to the absence of respective crystal structures, *Pa*LipA$_c$ and Lif:*Pa*LipA were built by homology modeling (Figures 2A-C). The models were assessed with our in-house model quality assessment program TopScore[17] and found to be 68 % correct for *Pa*LipA and 52 % correct for Lif. The correctness of the model is computed as the predicted global and local lDDT score[42], which compares all intra-molecular heavy-atom distances within two structures. If all distances deviate by more than 4 Å, the two structures are considered entirely different, and they are considered completely identical if all distances deviate by less than 0.5 Å. As the native structure is not known, the score is predicted by a deep neural network, which was trained on a large dataset of 660 protein targets totaling over 1.33×10^5 models and over 1.9×10^7 residues. It uses model quality predictions from different sources as input, including an agreement between features predicted from the sequence and measured in the model, such as secondary structure, solvent accessibility, and residue contacts.

First, we analyzed the MD simulations with respect to the average β-sheet propensity of residues 17-30 of *Pa*LipA because this secondary structure type is a characteristic feature of *Pa*LipA$_o$. For the Lif:*Pa*LipA complex, the likelihood of β-sheet formation is highest (96.5 ± 0.8%, mean ± SEM) (Figure 2D). In contrast, *Pa*LipA$_c$ showed a significantly lower β-sheet propensity of 41.1 ± 12%. This result indicates that Lif fosters the formation of the β-sheet structure. As expected, *Pa*LipA$_o$ exhibits a β-sheet propensity more similar to that of the Lif:*Pa*LipA complex (83.5 ± 3.5%), yet, the significantly smaller value suggests that *Pa*LipA$_o$ tends to move towards the closed conformation.

Next, we computed D_{COM} between H5 and H8 over the MD simulations to measure the opening and closing of the active site. Starting from *Pa*LipA$_c$, pronounced fluctuations of D_{COM} were observed that encompass both partially open lid conformations ($D_{COM} \approx 16$ Å) and more closed ones ($D_{COM} \approx 10$ Å) compared to the starting state ($D_{COM} = 13.3$ Å) (Figure 2E). A similar behavior was observed for *B. cepacia* lipase during MD simulations in water.[43] Starting with the Lif:*Pa*LipA complex, the probability density of partially open conformations ($D_{COM} \approx 16$ Å) was ~2-fold higher than for *Pa*LipA$_c$ (Figure 2E). This suggests that *Pa*LipA

in complex with Lif has a stronger tendency to move towards the partially open state than *Pa*LipA$_c$, although this tendency is obvious in only three trajectories out of six. As when starting from the closed conformation of *Pa*LipA, further closing of the lid was also observed during the MD simulations of the complex. Finally, starting from *Pa*LipA$_o$, the partially open state becomes most populated ($D_{COM} \approx 16$ Å), and even closed conformations ($D_{COM} \leq 13.3$ Å) were found (Figure 2E). Likewise, for *B. cepacia* lipase, an open-to-closed transition of the lid during MD simulations in water was found.[44]

To get an atomistic view on the further closed and partially open states observed in the above probability density distributions, we clustered the structures generated from the six MD trajectories for each system with respect to D_{COM}, using a threshold value of 4 Å. The two most populated clusters obtained respectively (Figure S1) were analyzed as to conformational changes in the lid domain. For *Pa*LipA$_c$ and the Lif:*Pa*LipA complex, the most populated clusters were dominated by structures with further closed active site ($D_{COM} < 13.3$ Å) (Figure S1A and S1C). In the second most populated clusters, in addition to lid movement towards larger D_{COM} values showing a partial opening of the active site, we also observed the formation of an additional α-helical structure for H5 of *Pa*LipA$_c$ and *Pa*LipA in complex with Lif (Figure S1B and S1D). For *Pa*LipA$_o$, the representative structures of the two most populated clusters ($D_{COM} \approx 16.2$ Å for the first and $D_{COM} \approx 13.9$ Å for the second, respectively (Figure S1E and S1F)) show a decrease and a bent in α-helix structure of H5 similar to Lif:*Pa*LipA (Figure S1D) when compared to the X-ray structure of *Pa*LipA$_o$.

To conclude, the lid of *Pa*LipA shows pronounced structural fluctuations on the μs time scale, reaching also a partially open state when starting from either a closed or open state. When starting from the closed state, reaching the partially open state is more favored for *Pa*LipA when bound to Lif.

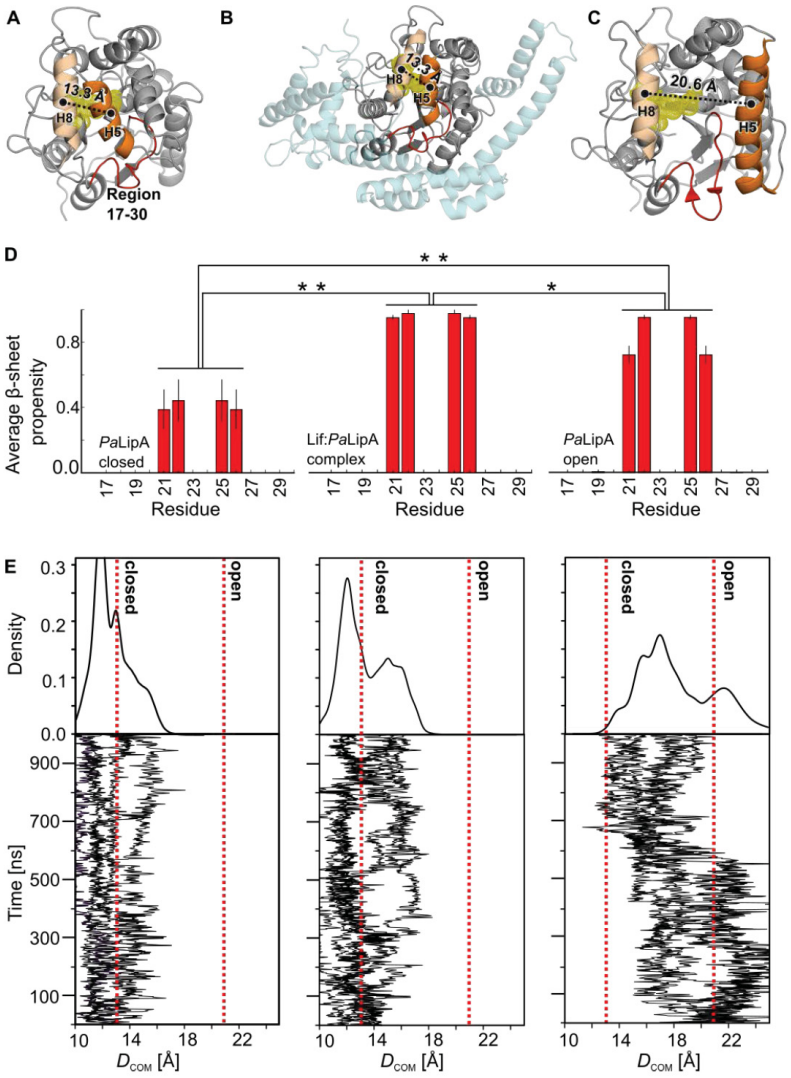


Figure 2. Structural dynamics of *Pa*LipA during unbiased MD simulations. (A) Structural model of *Pa*LipA in the closed conformation (*Pa*LipA$_c$) generated by homology modelling. Active site residues (catalytic triad residues S82, H251, and D229) are shown as yellow dots, which are covered by the lid domain (H5) (orange). H8 is highlighted in wheat. D_{COM} is represented by a black, dotted line. In *Pa*LipA$_c$, D_{COM} = 13.3 Å. Residues 17-30 are shown in red and do not exhibit the partial β-sheet structure, which is a characteristic feature of *Pa*LipA$_o$. **(B)** Homology model of the closed *Pa*LipA in complex with Lif (Lif:*Pa*LipA), Lif is represented in cyan, otherwise the representation is as in (A). **(C)** Crystal structure of *Pa*LipA$_o$ (PDB code: 1EX9), represented as in (A). H5 moved away from the active site, and D_{COM} = 20.6 Å (black dotted line). The residues 17-30 form a partial β-sheet structure (red). **(D)** Average per-residue β-sheet propensities of residues 17-30 starting from *Pa*LipA$_c$ (left),

Lif:*Pa*LipA (middle), and *Pa*LipA$_o$ (right). Error bars indicate SEM (eq. 1) and statistically significant differences of the averages calculated with the Student's *t*-test were indicated by "*" *p*-values < 0.05 and "**" *p*-values < 0.001. **(E)** D_{COM} over the simulation times of six MD trajectories each for the three systems listed in (D). Additionally, the probability densities are shown. Red dotted lines indicate the D_{COM} values of the open and closed *Pa*LipA states.

The open state of *Pa*LipA is more favorable when bound to Lif

To complement the unbiased MD simulations, we computed the potential of mean force (PMF) of the opening of the active site in free *Pa*LipA and complex with Lif (Lif:*Pa*LipA), applying umbrella sampling and using D_{COM} (Figure 2A) as a reaction coordinate. The PMF computations were performed for a plausible transition path of H5 obtained from unbiased MD simulations of *Pa*LipA$_o$ (see Methods section for details). Approximately Gaussian-shaped frequency distributions were obtained for each reference point along the reaction coordinate, with well overlapping windows (Figure S2). Such distributions are a prerequisite for the successful application of WHAM to compute a PMF.[26] Repeating the PMF computations for parts of the simulation time demonstrates that, for both systems, the PMFs are converged after at most 550 ns of simulation time per window (maximal difference between any two PMFs after a simulation time of 550 ns: 0.1 kcal mol^{-1}) (Figure S3). Usually, high configurational entropy results in the delayed convergence of the PMF.[45] Along these lines, our unbiased MD simulations reveal that the lid of *Pa*LipA fluctuates markedly on the µs time scale (Figure 2E). Furthermore, during lid opening, helix formation in the region of residues 125 – 135 takes place, as also indicated from the comparison of the open and closed structures (Figure 1A and 1C). Helix formation occurs on the time scale of hundreds of ns.[46,47] Both effects likely contribute that at most 550 ns per window are required to achieve converged PMFs. For comparison, the PMF values at the smallest D_{COM} sampled (11.6 Å) were set to zero in both cases (Figure 3A).

Although in both cases the configurational free energy increases with increasing D_{COM}, the PMFs differ in their global shape: The PMF of the Lif:*Pa*LipA complex increases more moderately than that of *Pa*LipA and shows broader local minima (Figure 3A). In more detail, the global energy minima (state I) for both free *Pa*LipA and the Lif:*Pa*LipA complex are found for the closed state (D_{COM} = 12.4 Å and 13.2 Å, respectively, $\Delta G \approx 0$ kcal mol^{-1} with respect to D_{COM} = 11.6 Å). At $D_{COM} \approx$ 14.8 Å, both PMFs have a local minimum (state II) of similar height ($\Delta G \approx 1$ kcal mol^{-1}). The corresponding energy well of the Lif:*Pa*LipA complex is extended until $D_{COM} \approx$ 16 Å. In contrast, the PMF for *Pa*LipA rises steeply immediately

following the local minimum. This finding coincides with a higher population of partially open structures found for Lif:*Pa*LipA in the unbiased simulations. Finally, flat PMF regions are found for both systems at $D_{COM} \approx 20.6$ Å (state III), but the configurational free energies with respect to the global minimum differ (Lif:*Pa*LipA: $\Delta G \approx 2.9$ kcal mol^{-1}, *Pa*LipA: $\Delta G \approx 4.6$ kcal mol^{-1}).

Furthermore, we computed the average β-sheet propensity of residues 17-30 of *Pa*LipA with and without Lif over the reweighted (unbiased) (Figure S4; see Methods section and ref. [28]) configurations from umbrella sampling for states I-III, respectively (Figure 3B). At the global minimum (state I), the β-sheet propensity averaged over windows 1 and 2 is significantly lower for *Pa*LipA (~21 ± 6%) than in state II, averaged over windows 3 and 4, and state III, averaged over windows 9 and 10 (~ 60 ± 4% and ~ 63 ± 3%, respectively). In contrast to *Pa*LipA, in Lif:*Pa*LipA, state I (~ 94 ± 0.8%) has a similar β-sheet propensity as state II and state III (~ 75 ± 2% and 84 ± 1%, respectively). The difference between the average β-sheet propensities of states I of *Pa*LipA and Lif:*Pa*LipA is highly statistically significant ($p < 0.001$). Similarly, for states II and III, the respective average β-sheet propensities of Lif:*Pa*LipA are significantly higher than those of *Pa*LipA ($p < 0.05$ for both states).

To conclude, the PMF computations reveal that the open state of *Pa*LipA is disfavored compared to the closed state but that in Lif:*Pa*LipA the open state is 1.7 kcal mol^{-1} more favorable than in *Pa*LipA. Furthermore, according to unbiased configurations from the umbrella sampling simulations, binding to Lif significantly favors the formation of the β-sheet in the region of residues 17-30, and this effect is most pronounced in the state I (~73 fold increase in the propensity).

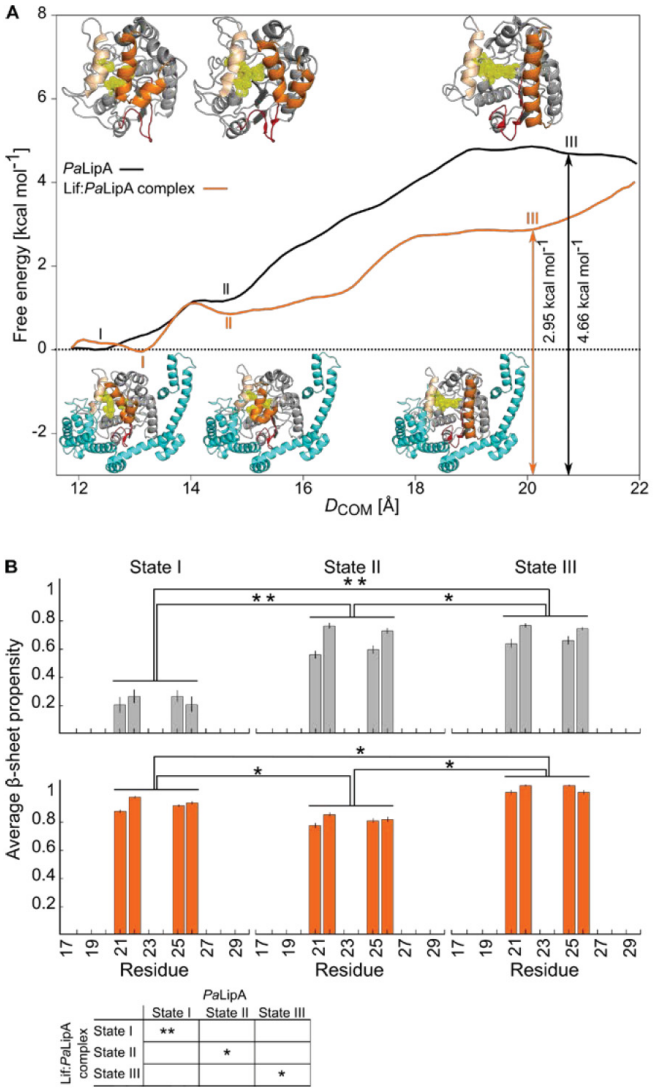


Figure 3. PMF computation of the active site opening in *Pa*LipA and the Lif:*Pa*LipA **complex and average β-sheet propensities of residues 17-30 of *Pa*LipA with and without Lif for three states identified in the PMFs. (A)** Configurational free energies of active site opening of *Pa*LipA as a function of D_{COM} used as a reaction coordinate for free *Pa*LipA (black) and the Lif:*Pa*LipA complex (orange). The standard deviation for all data points is < 0.002 kcal mol^{-1} computed by bootstrap analysis. Roman numbers indicate the identified states. Representative structures for states I-III are shown as cartoons for *Pa*LipA (top) and Lif:*Pa*LipA complex (bottom), respectively. The PMF values at D_{COM} = 11.6 Å were set to zero, respectively. **(B)** Per-residue averaged β-sheet propensity for residues 17-30 of *Pa*LipA, calculated across the umbrella sampling windows corresponding to the states I-III as described in (A), using reweighted (unbiased) configurations for *Pa*LipA (top) and the

Lif:*Pa*LipA complex (bottom). The table at the bottom displays results from comparing β-sheet propensities between *Pa*LipA and Lif:*Pa*LipA. Error bars indicate the SEM (eq. 1) and asterisks statistically significant differences (see Methods section for definition).

*Pa*LipA released from Lif loses its lipolytic activity over time under *in vitro* conditions

The unbiased MD simulations and the PMF computations reveal that *Pa*LipA$_o$ tends to move to an at most partially open state and that the open state is energetically disfavored with respect to the closed one, respectively. Although previous computations on related systems yielded similar results[44], our results are unexpected because, in a cellular context, secreted *Pa*LipA remains active and stable as indicated by the harsh conditions required for its denaturation.[14,48,49] In order to validate our computations, we thus performed biochemical experiments to probe if *Pa*LipA activity decreases under *in vitro* conditions similar to our simulations when the lipase is released from Lif.

To do so, a catalytically inactive *Pa*LipA variant, in which amino acid S82 of the catalytic triad is mutated to alanine, was used in addition to wild type *Pa*LipA. Purified *Pa*LipA$_{S82A}$ was renatured and used for complex formation with Lif at 1 µM concentration. According to the results of thermal unfolding experiments carried out with differential scanning fluorimetry (DSF), *Pa*LipA$_{S82A}$ forms a complex with Lif that has stability similar to that of *Pa*LipA with Lif (Figure 4A). This result is concordant with the fact that S82 is buried within *Pa*LipA and does not participate in interactions with Lif. At the used concentrations, the amount of free *Pa*LipA$_{S82A}$ or *Pa*LipA and Lif should be negligible because of the high binding affinity of Lif:*Pa*LipA (K_d = 5nM).

As expected, no catalytic activity is found for the Lif:*Pa*LipA$_{S82A}$ complex, in contrast to the Lif: *Pa*LipA complex (Figure 4B). After dilution of the Lif:*Pa*LipA$_{S82A}$ complex to 1 nM, renatured *Pa*LipA was added in excess at a concentration of 100 nM, followed by 3 h incubation. The addition of *Pa*LipA to the Lif:*Pa*LipA$_{S82A}$ complex restored activity to ~90 % of that of Lif:*Pa*LipA (Figure 4B), indicating that *Pa*LipA replaces *Pa*LipA$_{S82A}$ and then becomes activated by Lif.

Finally, we performed a complementary experiment in which 50 nM Lif:*Pa*LipA complex was supplemented with 50 nM or 100 nM *Pa*LipA$_{S82A}$, and with buffer as control, followed by

determination of the catalytic activity over time (Figure 4C). After 145 min, the activity level decreased by about 15 % and 35 % in the presence of 50 nM and 100 nM *Pa*LipA$_{S82A}$, respectively, which indicates the replacement of *Pa*LipA by *Pa*LipA$_{S82A}$ and the subsequent loss of catalytic activity of free *Pa*LipA in a *Pa*LipA$_{S82A}$ concentration-dependent manner. Addition of 100 nM Lif after 150 min restored catalytic activity, demonstrating that replaced *Pa*LipA can be re-activated by Lif.

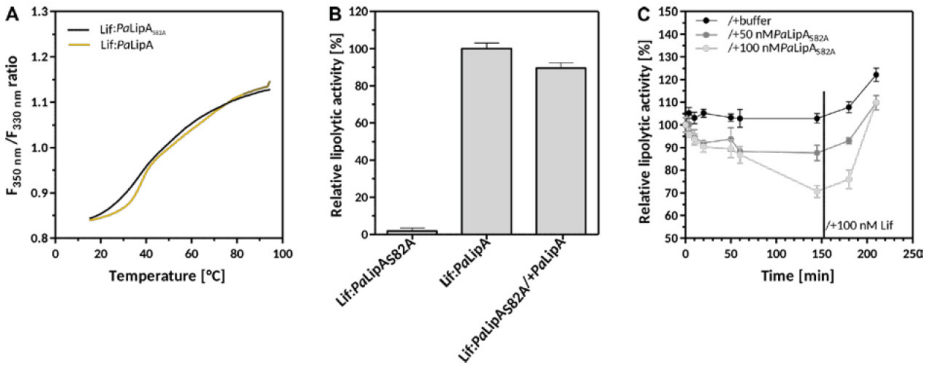


Figure 4. Dynamics of Lif:*Pa*LipA complex formation and *Pa*LipA activation. (A) DSF melting curves of the Lif:*Pa*LipA$_{S82A}$ and Lif:*Pa*LipA complexes at 1 µM concentrations. **(B)** The catalytic activity of Lif:*Pa*LipA$_{S82A}$ and Lif:*Pa*LipA (set to 100%) in comparison to Lif:*Pa*LipA$_{S82A}$ in the presence of *Pa*LipA (Lif:*Pa*LipA$_{S82A}$/+*Pa*LipA), which restores catalytic activity in the latter case. **(C)** The catalytic activity of Lif:*Pa*LipA over time in the absence (/+buffer, set to 100%) and presence of *Pa*LipA$_{S82A}$. The presence of *Pa*LipA$_{S82A}$ reduces catalytic activity in a concentration-dependent manner. The activity can be restored by addition of Lif (vertical line).

The activity decrease due to the addition of *Pa*LipA$_{S82A}$ in the first step of the experiment was lower than expected. According to the ratios of *Pa*LipA and *Pa*LipA$_{S82A}$, the expected activity decrease is 50% and 67% for the samples with 50 nM and 100 nM *Pa*LipA$_{S82A}$, respectively. The discrepancy is likely caused by incomplete complex formation at the start of the experiment and unfinished lipase exchange after 145 min. The fact that the Lif:*Pa*LipA control showed an increase in activity upon Lif addition supports the former point as does the lack of a plateau around 145 min in the case of 100 nM *Pa*LipA$_{S82A}$ and the need to incubate Lif with *Pa*LipA overnight to achieve maximal activation the latter.

To conclude, these *in vitro* experiments demonstrate that *Pa*LipA released from Lif loses its lipolytic activity over time and that the activity can be rescued by the addition of Lif. The results suggest that the closed-to-open transition of *Pa*LipA is a reversible process and that Lif

is required for the conformational transition of *Pa*LipA to the open state as well as to stabilize *Pa*LipA in the open conformation under *in vitro* conditions.

Lif binding affects the structural stability of key regions of *Pa*LipA involved in the opening of the active site

Our results suggest that Lif facilitates the opening of the active site in *Pa*LipA and stabilizes the partial β-sheet structure in the region of residues 17-30. To understand the underlying mechanism how Lif binding influences the active site opening in *Pa*LipA, we analyzed changes in the structural rigidity of *Pa*LipA upon mutating residues of Lif that interact with *Pa*LipA using an ensemble- and rigidity-theory based perturbation approach[36] integrated into the CNA approach.[37] Initially, we identified interactions between Lif and H5 as well as residues 17-30 of *Pa*LipA based on the Cα-Cα distance matrix averaged over the six unbiased MD simulations of the Lif:*Pa*LipA complex (Figure 5A). In total, 13 residues of Lif were identified (195-203, 213, 217-220) that are in direct contact with the region of residues 17-30 of *Pa*LipA (Figure 5B). By contrast, no residues of Lif were identified that interact with H5.

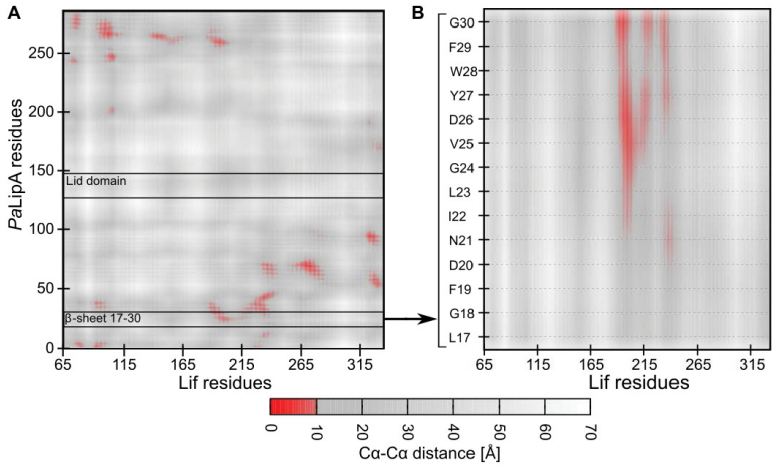


Figure 5. Lif residues interacting with *Pa*LipA. (A) Average Cα-Cα distance matrix calculated for the Lif:*Pa*LipA complex over six unbiased MD simulations of 1 μs length each. Residue pairs with a Cα-Cα distance < 10 Å are colored in red (see color scale) and considered in direct contact. Regions of H5 and residues 17-30 in *Pa*LipA are indicated by black lines. The SEM is < 0.1 Å in all cases. **(B)** Close-up of the Cα-Cα distance matrix for residues 17-30 in *Pa*LipA. Color code as in panel A.

To probe a potential influence of Lif binding on *Pa*LipA stability, first, a conformational ensemble of the Lif:*Pa*LipA complex was generated from the above shown MD simulations, constituting the ground state (see the Methods section for details). A perturbed state of the Lif:*Pa*LipA complex was then generated by removing the side chain of a Lif residue except the C_β atom, mimicking a substitution to alanine, but keeping the structures of Lif and *Pa*LipA unchanged otherwise. This perturbation was carried out separately for each of the 13 above mentioned residues. The changes are quantified as residue-wise free energy $\Delta G_{i,\mathrm{CNA}}$ (eq. 2), a measure for structural stability.[37] By definition, a change of the biomolecular conformation between the ground and perturbed states is excluded in our approach. Therefore, any observed changes in the biomolecular rigidity and flexibility must arise solely from local changes in the network topology that are due to the uncoupling of the residue side chain.[37] This procedure has been applied successfully before [37,50-52] and resembles a free energy decomposition scheme as non-perturbing alternative for (computational) alanine scanning mutagenesis.[53]

Of the 13 residues tested, F195, R199, R203, D218, and R219 showed the largest effect on the structural stability of *Pa*LipA (Figure 6, for effects on the structural stability of Lif see Figure S5, results for residues showing no effect are summarized in Figure S6). Upon perturbation of residue F195, the changes in $\Delta G_{i,\mathrm{CNA}}$ were largest for *Pa*LipA residues 15-45, which form the oxyanion hole and the cleft of the active site, whereas residue R199 affects the stability of residues 15-45 and in addition residues 255-268, which constitute the loop stabilizing the catalytic triad residue H251 (Figures 6A and 6B). Upon perturbation of residue R203, in addition to residues 15- 45, residues 142-144, which constitute the neighboring loop at the C-terminus of H5 also showed substantial changes in $\Delta G_{i,\mathrm{CNA}}$ (Figure 6C). By contrast, residues D218 and R219 specifically affected the stability of region 17-30 of *Pa*LipA (Figures 6D and 6E). Notably, all affected residues belong to the substrate binding site of *Pa*LipA which undergoes conformational rearrangements during activation. Finally, these perturbations also affect the stability of a number of the neighboring residues in Lif itself (Figure S5).

To conclude, the perturbation analysis reveals that certain Lif residues that directly interact with *Pa*LipA lead to a long-range impact on the structural stability of *Pa*LipA regions (residues 142-144, 255-268 and 15-45) in the vicinity of *Pa*LipA's active site. In particular, the stability of the partial β-sheet structure of residues 17-30 is affected, which forms upon opening of the lid domain.

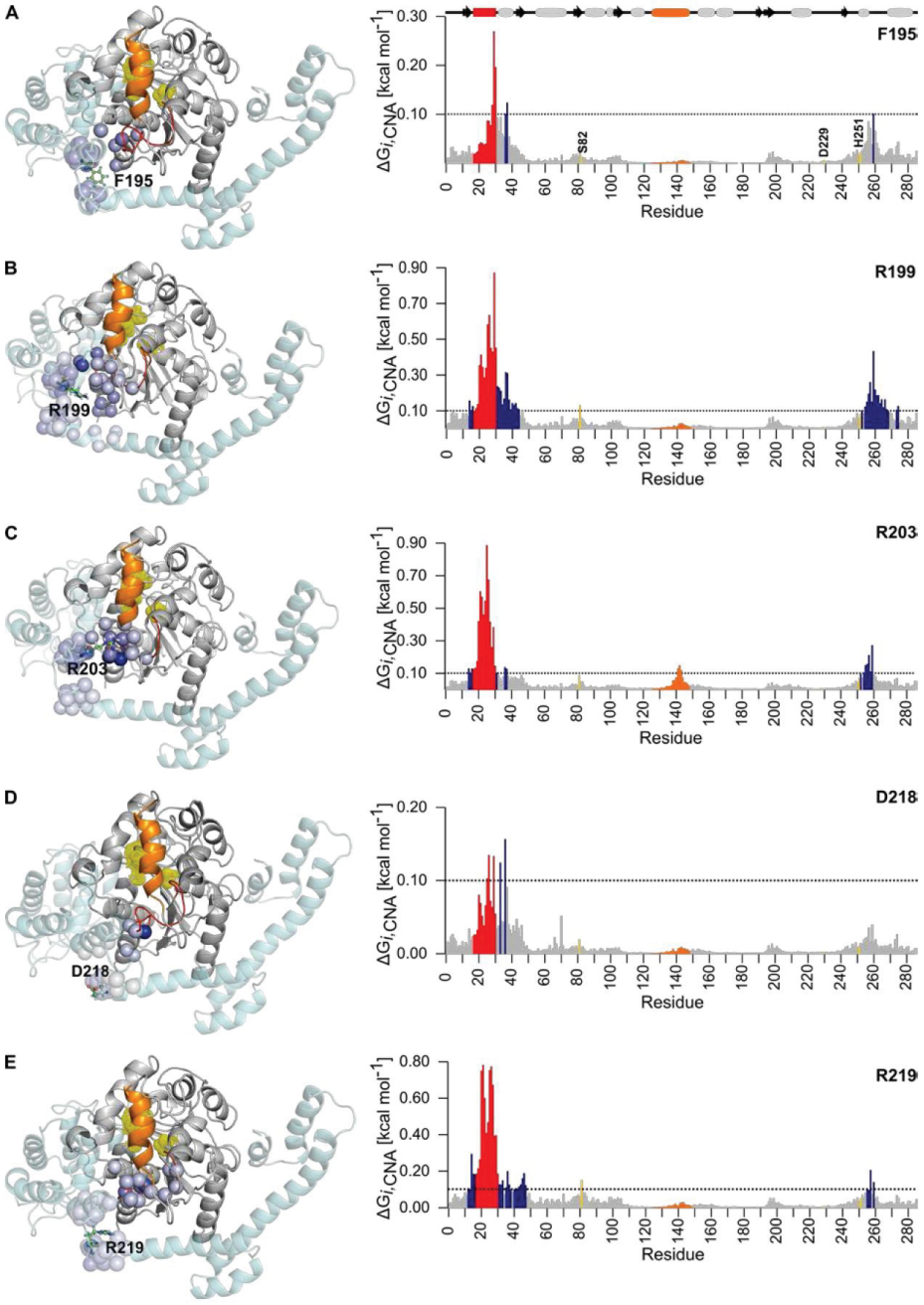


Figure 6. Potential influence of Lif residues interacting with *Pa*LipA on the structural stability of *Pa*LipA. A perturbation approach implemented in CNA was applied on the ensemble of structures of Lif:*Pa*LipA generated

by six unbiased MD simulations. **(A)** Left: Residues with $\Delta G_{i,\mathrm{CNA}}$ above the threshold of 0.1 kcal mol^{-1} are depicted as spheres on the Lif:*Pa*LipA complex structure. Blue colors reflect predicted $\Delta G_{i,\mathrm{CNA}}$ values, the larger the value, the darker is the color. The perturbed residue F195 of Lif (green, ball-and-stick representation) influences the stability of residues 17-30 of *Pa*LipA (red). H5 (orange) is shown in closed conformation occluding the binding site (yellow). Right: The histogram shows the per-residue $\Delta G_{i,\mathrm{CNA}}$ for *Pa*LipA (see Figure S5 for $\Delta G_{i,\mathrm{CNA}}$ of Lif). The dashed line indicates the threshold value of 0.1 kcal mol^{-1} above which residues are considered affected in terms of their structural stability. Residues forming H5, the catalytic triad, and region 17-30 are highlighted in orange, yellow, and red, respectively. Other residues with $\Delta G_{i,CNA}$ above the threshold are highlighted in blue. **(B)** As in panel A for the perturbation of residue R199 of Lif. **(C)** As in panel A for the perturbation of residue R203 of Lif. **(D)** As in panel A for the perturbation of residue D218 of Lif. **(E)** As in panel A for the perturbation of R219 of Lif. The standard error of the mean is < 0.05 kcal mol^{-1} for all residues in all cases.

Discussion

In this study, we have shown by molecular simulations at the atomistic level that the steric chaperone Lif catalyzes the activation process of *Pa*LipA by structurally stabilizing an intermediate *Pa*LipA conformation, particularly a β-sheet in the region of residues 17-30, such that the opening of the lid domain is facilitated. This opening allows substrate access to *Pa*LipA's catalytic site. Our study was motivated by previous experimental work that showed that the homologous *Bg*LipA in the absence of its foldase adopts a near-native conformation, which is enzymatically inactive, however.[10] Addition of the foldase results in lipase activity in solution.[10] In the Lif:*Bg*LipA complex crystal structure, *Bg*LipA shows a partial β-sheet formation in the region of residues 17-30,[13] which has also been observed in *Pa*LipA$_o$[1] but not in closed *Bg*LipA.[14] Notably, lipase activity was also found in the crystals of the Lif:*Bg*LipA complex,[13] concordant with helix 5 of *Bg*LipA having sufficient space to move in the crystal lattice and, thus, being able to switch to the open state.[9,54]

Our result is supported by three complementary computational approaches and *in vitro* biochemical experiments. First, we performed six independent, unbiased, microsecond-long MD simulations at the atomistic level in explicit solvent starting from (free) *Pa*LipA$_c$ and *Pa*LipA$_c$ in complex with Lif (Lif:*Pa*LipA). These simulations revealed that the lid of *Pa*LipA shows pronounced structural fluctuations on the μs time scale and also reaches a partially open state when starting from either a closed or open state. Yet, when starting from the closed state, the partially open state is reached more frequently if *Pa*LipA is bound to Lif than when it is free. To our knowledge, the length of our MD simulations surpasses comparable previous ones

on *Pa*LipA by at least 800 ns,[55-57] whereas no MD simulations have been reported for Lif:*Pa*LipA complex so far. For the MD simulations, we used established parameterizations for the solvent,[21] and proteins,[19] which we had also applied successfully in other simulations on soluble proteins.[29,38,58,59] Furthermore, the impact of force field deficiencies on our results is expected to be small due to cancellation of errors when comparatively assessing simulation results for *Pa*LipA in unbound or bound state, or started from different conformations. While an experimental structure was available for $PaLipA_o$,[1] homology models were used as starting structures for $PaLipA_c$ and Lif:*Pa*LipA. Note that, as the partial β-sheet structure is absent in the closed *Bg*LipA structure, our homology models of $PaLipA_c$ and Lif:*Pa*LipA do not have a partial β-sheet structure in the region of residues 17-30 either. Still, a much higher β-sheet propensity is found in that region in MD simulations of the Lif:*Pa*LipA complex than for $PaLipA_c$, concordant with the presence of such a β-sheet in the crystal structure of Lif:*Bg*LipA and the lack of it in the crystal structure of closed *Bg*LipA. Concomitantly, MD simulations started from $PaLipA_o$ exhibited a β-sheet propensity more similar to that of Lif:*Pa*LipA, although the smaller values suggested that $PaLipA_o$ tends to move towards the closed conformation. Apparently, binding to Lif fosters β-sheet formation in *Pa*LipA in the region of residues 17-30.

As a second, independent approach, we investigated the energetics of active site opening in free *Pa*LipA and in the complex with Lif (Lif:*Pa*LipA) by umbrella sampling simulations followed by PMF computations, using established protocols successfully applied previously by us[58,59] and D_{COM} as an intuitive reaction coordinate previously applied on a similar system.[58] To our knowledge, the energetics of active site opening in *Pa*LipA has not been investigated by computational means before. The PMF computations reveal that the open state of *Pa*LipA is disfavored compared to the closed state but that in Lif:*Pa*LipA the open state is 1.7 kcal mol^{-1} more favorable than in *Pa*LipA. Both findings are in agreement with results from unbiased MD simulations (see above), demonstrating internal consistency of our findings. The former finding is also in agreement with our *in vitro* experiments according to which *Pa*LipA set free from Lif loses its lipolytic activity over time. Finally, evaluating the β-sheet propensity of the region of residues 17-30 on reweighted configurations from the umbrella sampling simulations confirmed that binding to Lif significantly favors β-sheet formation in that region, particularly in the closed state, again demonstrating internal consistency with respect to results from unbiased MD simulations.

Third, we applied a rigidity theory- and ensemble-based perturbation approach for analyzing biomolecular rigidity and flexibility[36,37] successfully used previously by us[38] to scrutinize the mechanism of how Lif binding influences the active site opening in *Pa*LipA. The results revealed that five out of 13 Lif residues forming contacts with the region of residues 17-30 of *Pa*LipA contribute to the structural stability of the binding site in a long-range manner, in particular the region formed by residues 17-30, the neighboring loop of the lid/H5 (residues 142-144), and the loop (residues 255-268) stabilizing H251, one of the residues of the catalytic triad. Considering that tertiary interactions can stabilize β-sheet formation,[60-62] these results can explain why a higher β-sheet propensity in the region of residues 17-30 is found when *Pa*LipA is bound to Lif.

A surprising and so far not fully understood aspect of our study is that the open state of *Pa*LipA is unstable compared to the closed one according to our computational and *in vitro* results (Figure 7). This finding does not contradict results on α-lytic protease and subtilisin, which need steric chaperones to reach their active state and whose active states are less than or only marginally more stable than the inactive intermediate states.[12,63-65] Yet, it is at variance with the fact that *Pa*LipA secreted to the extracellular medium remains active.[14,48,49] At present, we can only speculate that further interactions of *Pa*LipA with the Xcp secretion machinery and/or components of the extracellular matrix contribute to the remaining activity, e.g., by increasing the energy barrier between the open and closed states of *Pa*LipA, which is almost absent in our free energy profiles.

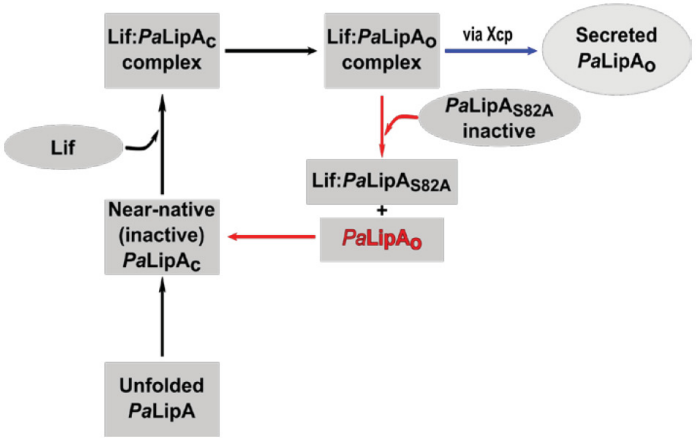


Figure 7. Proposed scheme for *Pa*LipA activation and secretion.[7,8,10] Unfolded *Pa*LipA folds to a near-native inactive state in the absence of Lif (*Pa*LipA$_C$). This *Pa*LipA$_C$ binds to Lif (Lif:*Pa*LipA$_C$) and undergoes

conformational changes towards the open state (Lif: *Pa*LipA$_o$). After opening, *Pa*LipA$_o$ is released from the Lif:*Pa*LipA$_o$ complex and *in vivo* secreted to the extracellular medium via the Xcp machinery (blue arrow), although the exact mechanism of secretion is still unknown. Results from MD simulations, PMF computations, and *in vitro* experiments (red arrows) obtained in this study indicate that *Pa*LipA$_o$ released from the complex (red) by addition of the inactive variant (*Pa*LipA$_{S82A}$) can fold back to inactive *Pa*LipA$_c$. Black arrows indicate processes occurring *in vitro* and *in vivo*. Our results thus suggest that Lif is required to facilitate the closed-to-open transition of *Pa*LipA$_c$ as well as to stabilize *Pa*LipA$_o$ under *in vitro* conditions until it is secreted to extracellular medium via the Xcp machinery.

In summary, our results shed light onto the molecular mechanism of a steric chaperone in that they provide an explanation how Lif directly catalyzes the folding process of *Pa*LipA by imprinting the essential steric (structural) information onto the target protein: Lif structurally stabilizes an intermediate *Pa*LipA conformation, particularly a β-sheet in the region of residues 17-30, such that the opening of *Pa*LipA's lid domain is facilitated.

Acknowledgments

This study was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through funding no. GO 1367/1-1 and INST 208/704-1 FUGG to HG and JA 448/8-1 to KEJ. Part of the work of FK and KEJ was funded by CRC 1208 (project A02). HG is grateful for computational support and infrastructure provided by the "Zentrum für Informations- und Medientechnologie" (ZIM) at the Heinrich Heine University Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) on the supercomputer JURECA at Jülich Supercomputing Centre (JSC) (user IDs: HKF7, HDD16; project IDs: 11237, 12732).

References

1. Nardini, M.; Lang, D. A.; Liebeton, K.; Jaeger, K.-E.; Dijkstra, B. W. J Biol Chem 2000.
2. Jaeger, K.-E.; Eggert, T. Curr Opin Biotechnol 2002, 13(4), 390-397.
3. Verger, R. Trends Biotechnol 1997, 15(1), 32-38.
4. Derewenda, Z. S. In Adv Protein Chem; Elsevier, 1994, p 1-52.
5. Jaeger, K.-E.; Reetz, M. T. Trends Biotechnol 1998, 16(9), 396-403.
6. Rosenau, F.; Jaeger, K.-E. Biochimie 2000, 82(11), 1023-1032.
7. Rosenau, F.; Tommassen, J.; Jaeger, K. E. ChemBioChem 2004, 5(2), 152-161.
8. Douzi, B.; Ball, G.; Cambillau, C.; Tegoni, M.; Voulhoux, R. J Biol Chem 2011, 286(47), 40792-40801.
9. El Khattabi, M.; Van Gelder, P.; Bitter, W.; Tommassen, J. J Biol Chem 2000, 275(35), 26885-26891.
10. Pauwels, K.; del Pino, M. M. S.; Feller, G.; Van Gelder, P. PLoS One 2012, 7(5), e36999.

11. Pauwels, K.; Van Molle, I.; Tommassen, J.; Van Gelder, P. Mol Microbiol 2007, 64(4), 917-922.
12. Sohl, J. L.; Jaswal, S. S.; Agard, D. A. Nature 1998, 395(6704), 817.
13. Pauwels, K.; Lustig, A.; Wyns, L.; Tommassen, J.; Savvides, S. N.; Van Gelder, P. Nat Struct Mol Biol 2006, 13(4), 374-375.
14. Noble, M.; Cleasby, A.; Johnson, L.; Egmond, M.; Frenken, L. FEBS Lett 1993, 331(1-2), 123-128.
15. Kelley, L. A.; Sternberg, M. J. E. Nature Protocols 2009, 4, 363.
16. Guex, N.; Peitsch, M. C. Electrophoresis 1997, 18(15), 2714-2723.
17. Mulnaes, D.; Gohlke, H. Journal of chemical theory and computation 2018, 14(11), 6117-6126.
18. Case, D.; Babin, V.; Berryman, J.; Betz, R.; Cai, Q.; Cerutti, D.; Darden, T.; Duke, R.; Gohlke, H.; Goetz, A. AMBER14 San Francisco: University of California 2014.
19. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. J Chem Theory Comput 2015, 11(8), 3696-3713.
20. Ciglia, E.; Vergin, J.; Reimann, S.; Smits, S. H.; Schmitt, L.; Groth, G.; Gohlke, H. PLOS one 2014, 9(4), e96031.
21. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. J Chem Phys 1983, 79(2), 926-935.
22. Darden, T.; York, D.; Pedersen, L. J Chem Phys 1993, 98(12), 10089-10092.
23. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. J Comput Phys 1977, 23(3), 327-341.
24. Berendsen, H. J.; Postma, J. v.; van Gunsteren, W. F.; DiNola, A.; Haak, J. The Journal of chemical physics 1984, 81(8), 3684-3690.
25. Simmerling, C.; Strockbine, B.; Roitberg, A. E. J Am Chem Soc 2002, 124(38), 11258-11259.
26. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. J Comput Chem 1992, 13(8), 1011-1021.
27. Roe, D. R.; Cheatham, T. E. Journal of Chemical Theory and Computation 2013, 9(7), 3084-3095.
28. Moradi, M.; Enkavi, G.; Tajkhorshid, E. Nature communications 2015, 6, 8393.
29. Pagani, G.; Gohlke, H. Scientific reports 2018, 8(1), 5733.
30. Chodera, J. D. Journal of chemical theory and computation 2016, 12(4), 1799-1805.
31. Welch, B. L. Biometrika 1947, 34(1/2), 28-35.
32. Fox, J.; Andersen, R. Department of Sociology, McMaster University 2005, 2-4.
33. Li, M. Z.; Elledge, S. J. In Gene synthesis; Springer, 2012, p 51-59.
34. Hausmann, S.; Wilhelm, S.; Jaeger, K. E.; Rosenau, F. FEMS Microbiol Lett 2008, 282(1), 65-72.
35. Jaeger, K.-E.; Kovacic, F. In Pseudomonas Methods and Protocols; Springer, 2014, p 111-134.
36. Pfleger, C.; Minges, A.; Boehm, M.; McClendon, C. L.; Torella, R.; Gohlke, H. Journal of Chemical Theory and Computation 2017, 13(12), 6343-6357.
37. Pfleger, C.; Rathi, P. C.; Klein, D. L.; Radestock, S.; Gohlke, H. Journal of Chemical Information and Modeling 2013, 53(4), 1007-1015.
38. Milić, D.; Dick, M.; Mulnaes, D.; Pfleger, C.; Kinnen, A.; Gohlke, H.; Groth, G. Scientific Reports 2018, 8(1), 3890.
39. Hermans, S. M.; Pfleger, C.; Nutschel, C.; Hanke, C. A.; Gohlke, H. Wiley Interdisciplinary Reviews: Computational Molecular Science 2017, 7(4), e1311.
40. Pfleger, C.; Gohlke, H. Structure 2013, 21(10), 1725-1734.
41. Pfleger, C.; Radestock, S.; Schmidt, E.; Gohlke, H. J Comput Chem 2013, 34(3), 220-233.

42. Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T. Bioinformatics 2013, 29(21), 2722-2728.
43. Trodler, P.; Schmid, R. D.; Pleiss, J. BMC Struct Biol 2009, 9(1), 38.
44. Lee, H. S.; Oh, Y.; Kim, M.-J.; Im, W. The Journal of Physical Chemistry B 2018, 122(47), 10659-10668.
45. Joshi, D. C.; Lin, J. H. J Comput Chem 2019, 40(17), 1652-1663.
46. Lin, M. M.; Mohammed, O. F.; Jas, G. S.; Zewail, A. H. Proceedings of the National Academy of Sciences 2011, 108(40), 16622-16627.
47. De Sancho, D.; Best, R. B. J Am Chem Soc 2011, 133(17), 6809-6816.
48. El Khattabi, M.; Van Gelder, P.; Bitter, W.; Tommassen, J. J Mol Catal B: Enzym 2003, 22(5-6), 329-338.
49. Frenken, L.; Egmond, M. R.; Batenburg, A.; Bos, J. W.; Visser, C.; Verrips, C. T. Appl Environ Microbiol 1992, 58(12), 3787-3791.
50. Milic, D.; Dick, M.; Mulnaes, D.; Pfleger, C.; Kinnen, A.; Gohlke, H.; Groth, G. Sci Rep 2018, 8(1), 3890.
51. Rathi, P. C.; Fulton, A.; Jaeger, K.-E.; Gohlke, H. PLOS Comp Biol 2016, 12, e1004754.
52. Rathi, P. C.; Jaeger, K. E.; Gohlke, H. PLoS One 2015, 10(7), e0130289.
53. Gohlke, H.; Kiel, C.; Case, D. A. J Mol Biol 2003, 330, 891-913.
54. El Khattabi, M.; Ockhuijsen, C.; Bitter, W.; Jaeger, K.-E.; Tommassen, J. Molecular and General Genetics MGG 1999, 261(4-5), 770-776.
55. Cherukuvada, S. L.; Seshasayee, A. S. N.; Raghunathan, K.; Anishetty, S.; Pennathur, G. PLoS Comp Biol 2005, 1(3), e28.
56. Ferrario, V.; Ebert, C.; Knapic, L.; Fattor, D.; Basso, A.; Spizzo, P.; Gardossi, L. Adv Synth Catal 2011, 353(13), 2466-2480.
57. Johnson, Q. R.; Lindsay, R. J.; Nellas, R. B.; Shen, T. Proteins: Structure, Function, and Bioinformatics 2016, 84(6), 820-827.
58. Ciupka, D.; Gohlke, H. Scientific Reports 2017, 7(1), 8020.
59. Minges, A.; Ciupka, D.; Winkler, C.; Höppner, A.; Gohlke, H.; Groth, G. Scientific Reports 2017, 7, 45389.
60. Hutchinson, E. G.; Sessions, R. B.; Thornton, J. M.; Woolfson, D. N. Protein Sci 1998, 7(11), 2287-2300.
61. Merkel, J. S.; Sturtevant, J. M.; Regan, L. Structure 1999, 7(11), 1333-1343.
62. Wouters, M. A.; Curmi, P. M. G. Proteins: Structure, Function, and Bioinformatics 1995, 22(2), 119-131.
63. Bryan, P.; Wang, L.; Hoskins, J.; Ruvinov, S.; Strausberg, S.; Alexander, P.; Almog, O.; Gilliland, G.; Gallagher, T. Biochemistry 1995, 34(32), 10310-10318.
64. Jaswal, S. S.; Sohl, J. L.; Davis, J. H.; Agard, D. A. Nature 2002, 415(6869), 343.
65. Jaswal, S. S.; Truhlar, S. M.; Dill, K. A.; Agard, D. A. J Mol Biol 2005, 347(2), 355-366.

SUPPORTING INFORMATION

The membrane-integrated steric chaperone Lif facilitates active site opening of *Pseudomonas aeruginosa* lipase A

Neha Verma[1], Peter Dollinger[2], Filip Kovacic[2], Karl-Erich Jaeger[2,3], and Holger Gohlke[1,4] *

[1]Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

[2]Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, Forschungszentrum Jülich GmbH, 52426 Jülich, Germany

[3]Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52426 Jülich, Germany

[4]John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC) and Institute for Complex Systems - Structural Biochemistry (ICS-6), Forschungszentrum Jülich GmbH, 52426 Jülich, Germany

*Address: Universitätstr. 1, 40225 Düsseldorf, Germany.
Phone: (+49) 211 81 13662; Fax: (+49) 211 81 13847
E-mail: gohlke@uni-duesseldorf.de

Supplementary Figures

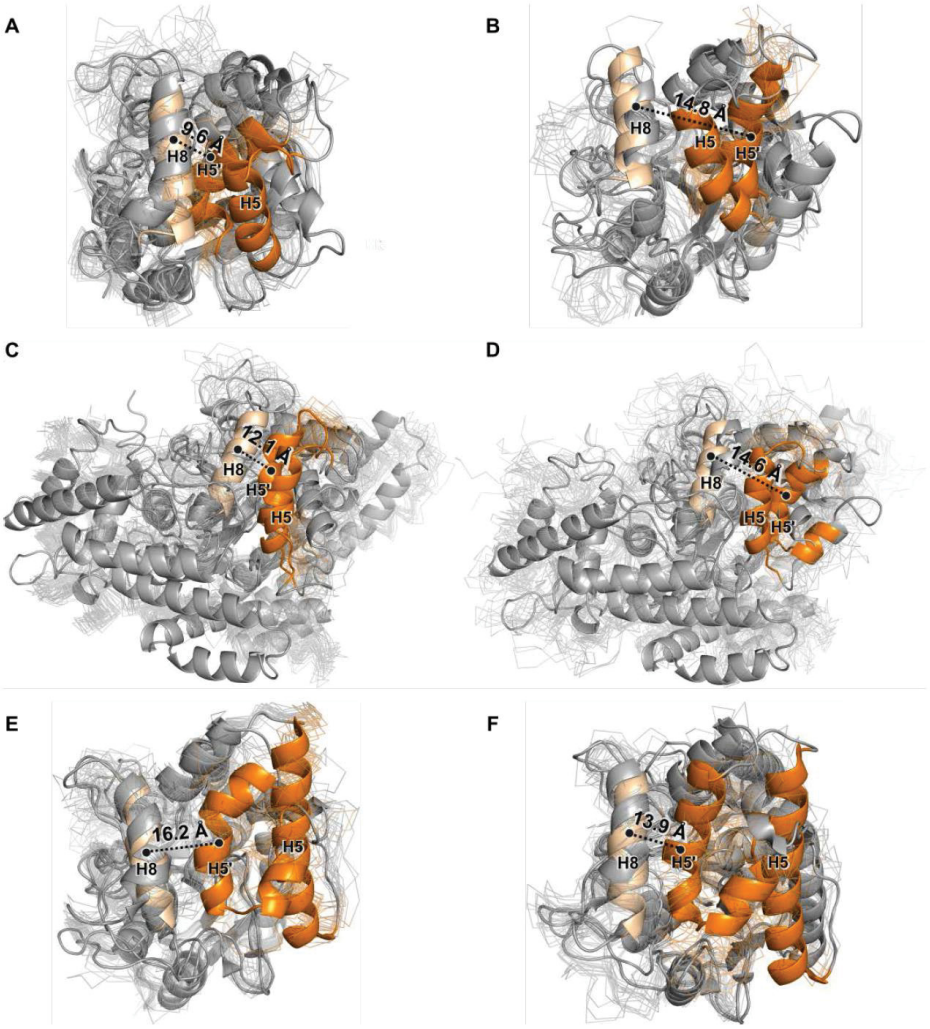


Figure S1. Structural clusters of *Pa*LipA, Lif:*Pa*LipA complex, and open *Pa*LipA. All structures obtained over the simulation times of six MD trajectories for the three systems were clustered with respect to D_{COM} between H5 and H8 with a cutoff of 4 Å. Two most populated clusters were identified for each system, with the representative structure shown as cartoon (lid domain H5' (orange)) aligned with the starting structure (lid domain H5 (orange) and H8 (wheat)), highlighting the conformational changes in lid domain. Further members of clusters are shown in ribbon representation. **(A)** Most populated cluster obtained for *Pa*LipA; the representative structure shows the further closing of the lid with D_{COM} = 9.6 Å. **(B)** Second most populated cluster obtained for *Pa*LipA, with D_{COM} = 14.8 Å for the representative structure indicating partial opening of the lid. **(C)** For the Lif:*Pa*LipA complex, the

representative structure of the most populated cluster shows further closing of the lid (D_{COM} = 12.1 Å) and is aligned with the starting homology model of the Lif:*Pa*LipA complex. **(D)** The second most populated cluster of the Lif:*Pa*LipA complex shows partial opening of the active site with D_{COM} = 14.6 Å for the representative structure. **(E)** The representative structure of the most populated cluster is aligned with the X-ray structure of open *Pa*LipA. D_{COM} = 16.2 Å indicates a partial closing of the lid for open *Pa*LipA. **(F)** The representative structure of the second most populated cluster with D_{COM} = 13.9 Å indicates a complete closure of the lid, starting from the open *Pa*LipA conformation.

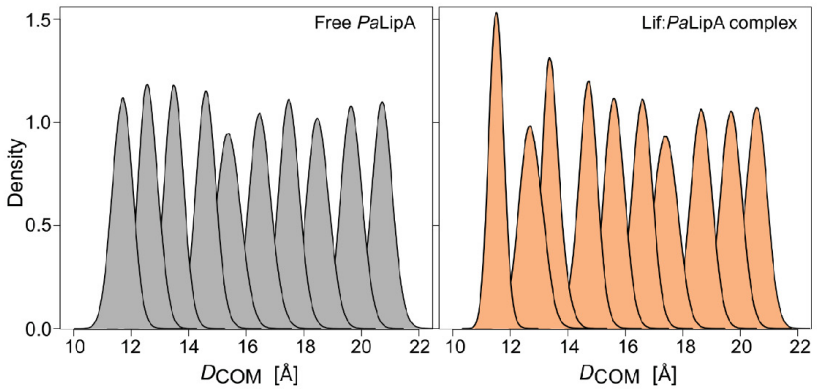


Figure S2. Overlap of umbrella sampling simulations of the lid opening in free *Pa*LipA (left) and in the Lif:*Pa*LipA complex (right). Frequency distributions of values of the reaction coordinate obtained from umbrella sampling simulations are shown.

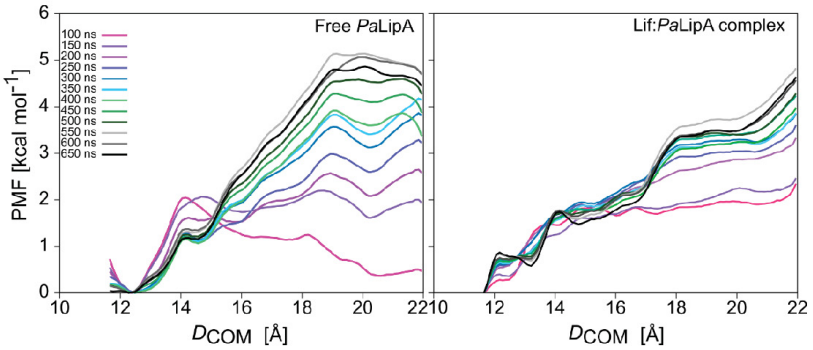


Figure S3. Convergence of the PMFs of the lid opening in free *Pa*LipA (left) and the Lif:*Pa*LipA complex (right). The development of the PMF profiles computed for the lid opening over the umbrella sampling time as indicated by the color scale at the left is shown.

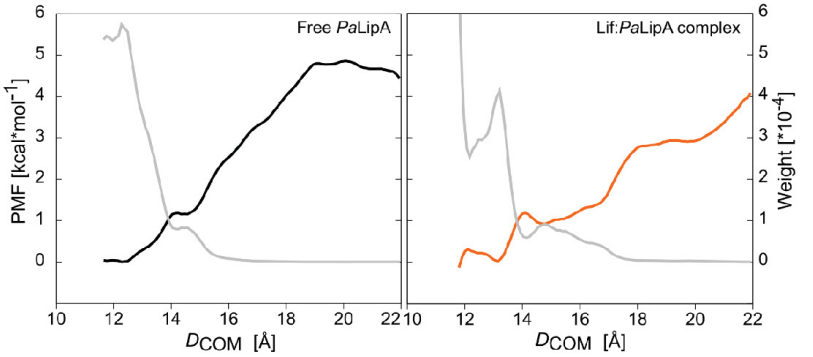


Figure S4. Reweighting of configurations obtained from umbrella sampling of active site opening in free *Pa*LipA (left) and the Lif:*Pa*LipA complex (right). The curves colored gray (right ordinate values) depict the weight w^t of a configuration t computed according to eqs. 7 and 8 from ref. [1], which is normalized by dividing it by the sum of all weights for each system. The respective PMF profile (black and orange; left ordinate values) is overlaid. The reweighting is performed considering the entire ensemble for each system between reaction coordinate values 11.6 to 20.6 Å.

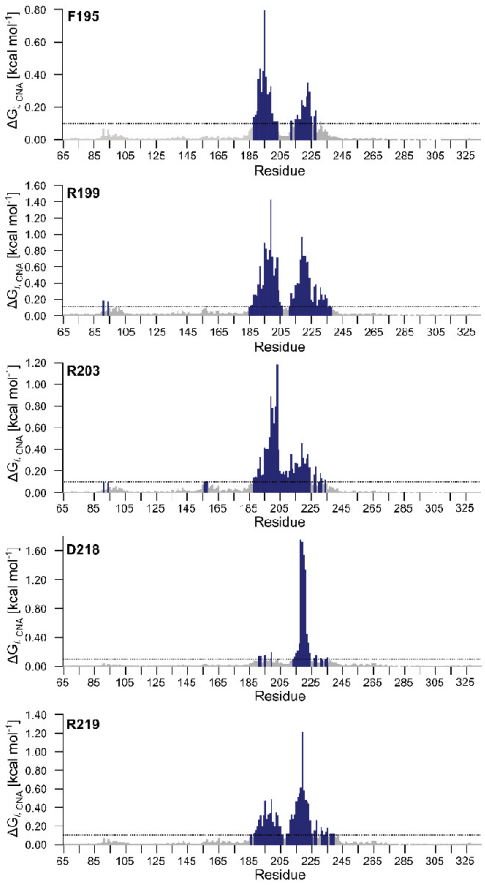


Figure S5. Per-residue $\Delta G_{i,CNA}$ for Lif calculated by CNA[2] on an ensemble of structures of the Lif:*Pa*LipA complex generated from six unbiased MD simulations. The dashed line at 0.1 kcal mol^{-1} indicates the threshold value, and residues with higher $\Delta G_{i,CNA}$ values are colored blue. The standard error of mean is < 0.05 kcal mol^{-1} for all residues. From top to bottom, $\Delta G_{i,CNA}$ for Lif are shown when perturbing the complex by removing constraints involving Lif residues F195, R199, R203, D218, and R219, respectively.

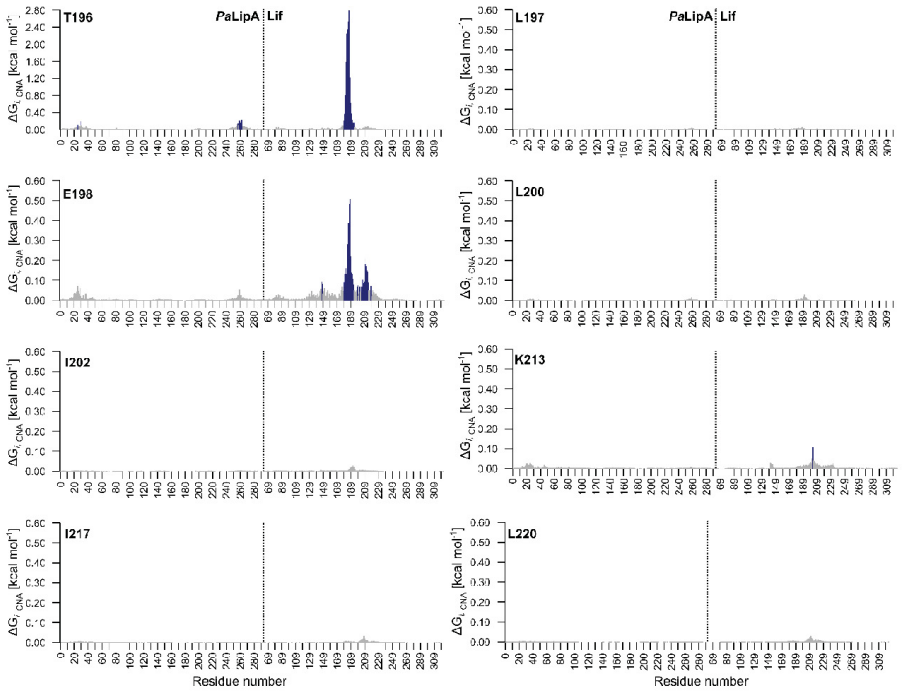


Figure S6. Per-residue $\Delta G_{i,CNA}$ for *Pa*LipA and Lif calculated by CNA on an ensemble of structures of the Lif:*Pa*LipA complex are shown upon perturbation of Lif residues 196-200, 202, 213, 217, and 220. The standard error of the mean is < 0.05 kcal mol^{-1} for all residues. The residues with $\Delta G_{i,CNA}$ value above the threshold value are shown in blue, and others are in gray. In all plots, the respective perturbations do not affect the stability of the *Pa*LipA residues above the threshold value of 0.1 kcal mol^{-1}.

Supplementary References

1. Moradi, M.; Enkavi, G.; Tajkhorshid, E. Nat. Commun. 2015, 6, 8393.
2. Pfleger, C.; Rathi, P. C.; Klein, D. L.; Radestock, S.; Gohlke, H. J. Chem. Inf. Model 2013, 53(4), 1007-1015.

CURRICULUM VITAE

NEHA VERMA

Date of Birth: 2nd November 1989

Birth Place: Bareilly, INDIA

EDUCATION:

Ph.D., Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany, Computational Pharmaceutical Chemistry and Molecular Bioinformatics (2014 - present)
Dissertation: "Energetics and dynamics of the lipase activation by the steric chaperone Lif in *Pseudomonas aeruginosa"*

MS, NIPER University, Mohali, India, Pharmacoinformatics (2012 - 2014)
Thesis: "Metabolic inhibition of intestinal UGTs by pharmaceutical excipients: a mechanistic study using molecular dynamics and *in vitro* studies"

BS, GBT University, India, Pharmacy (2008 - 2012)

RELEVANT TRAINING:

Markov State Modeling and Free Energy Calculation Workshop, Freie Universität, Berlin, Germany. 2016

Schrödinger Workshop, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. 2017

Good Scientific Practice and Optimizing writing Strategies for publishing research Workshop, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. 2016

TEACHING EXPERIENCE:

Heinrich-Heine-Universität, Dept. of Pharmacy, 2014 – 2019, laboratory instructor and teaching assistant for Pharmacy students within "quantitative analysis" practical course

SELECTED GRANTS, PRESENTATIONS AND PUBLICATIONS:

Selected grants

Grant for computing time on Supercomputers at JSC, 2016-2017, Conformational dynamics of the unbound lipase-specific foldase Lif

Grant for computing time on Supercomputers at JSC, 2017-2018, Analysis of the conformational changes during activation of lipase A by its foldase

Presentations and posters

Neha Verma
"Conformational dynamics of the steric chaperon MD simulations with MSM and FRET". Workshop on Markov modeling and Free Energy calculation, Freie Universität, Berlin, 2016

Neha Verma
"Characteristics of the unbound state of Lipase foldase by integrative modeling". Institute seminar, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany, 2017

Neha Verma
"Quantitative characterization of the conformational dynamics of the unbound lipase-specific foldase Lif by MD simulations and fluorescence spectroscopy". 62nd Annual Meeting of the Biophysical Society, San Francisco, USA, 2018

Neha Verma, Holger Gohlke
"Conformational dynamics of unbound of the unbound lipase-specific foldase Lif studied by MD simulations and fluorescence spectroscopy". NIC symposium, Jülich, Germany, 2018

Publications

N. W. Hieke, N. Verma, D. Schlütermann, N. Berleth, J. Deitersen, P. Böhler, F. Stuhldreier, W. Wu, S. Seggewiß, C. Peter, H. Gohlke, N. Mizushima, B. Stork. *Systematic analysis of ATG13 domain requirements for autophagy induction*. Autophagy, pages 1-58, 2017.

S. P. Tripathi, R. Prajapati, N. Verma, A. T. Sangamwar. *Predicting substrate selectivity between UGT1A9 and UGT1A10 using molecular modeling and molecular dynamics approach*. Molecular Simulations, Vol. 42, 2015.

In submission and preparation

Aldino Viegas, Peter Dollinger, Neha Verma, Jakub Kubiak, Thibault Viennet, Claus A.M. Seidel, Holger Gohlke, Manuel Etzkorn, Filip Kovacic, and Karl-Erich Jaeger. *Structural and dynamic insights revealing how lipase binding domain MD1 of Pseudomonas aeruginosa foldase affects lipase activation.*

Jakub Kubiak, Neha Verma, Peter Dollinger, Holger Gohlke, Claus A. M. Seidel, Filip Kovacic, Karl-Erich Jaeger. *Functional dynamics of a structurally minimalistic chaperone.*

Neha Verma, Peter Dollinger, Filip Kovacic, Karl-Erich Jaeger, and Holger Gohlke. *The membrane-integrated steric chaperone Lif facilitates active site opening of Pseudomonas aeruginosa lipase A*.